Contents

1 Manifolds and tensors 4
  1.1 Topological definitions ............................................. 4
  1.2 Manifolds .......................................................... 4
  1.3 Partitions of unity .................................................. 5
  1.4 Vectors .............................................................. 5
  1.5 Covectors .......................................................... 7
  1.6 Tensors ............................................................. 8
  1.7 Contraction of tensors .............................................. 9
  1.8 Non-coordinate bases ............................................... 10
  1.9 Abstract index notation ........................................... 10
  1.10 Tensor fields, and the commutator of vector fields .......... 11
  1.11 The metric tensor ................................................ 12
  1.12 Exercises ......................................................... 15

2 Maps of manifolds, integral curves and the Lie derivative 16
  2.1 Motivation: derivatives of vector fields ....................... 16
  2.2 Maps of manifolds .................................................. 16
  2.3 Integral curves and 1-parameter families of maps ............. 18
  2.4 Lie derivative of a function ..................................... 18
  2.5 Lie derivative of a vector field ................................ 19
  2.6 Lie derivative of a covector field ................................ 20
  2.7 Lie derivative of a tensor field of arbitrary rank .......... 20
  2.8 Coordinates adapted to a vector field ......................... 20
  2.9 Submanifolds ....................................................... 21
  2.10 Application: axisymmetry ........................................ 22
  2.11 Exercises ........................................................ 23

3 Linear connections and curvature 25
  3.1 Linear connections ................................................ 25
  3.2 Parallel transport and geodesics ................................. 26
  3.3 The exponential map .............................................. 28
  3.4 Torsion ............................................................... 28
  3.5 Curvature ........................................................... 29
  3.6 Flat connections .................................................. 31
  3.7 Geodesic deviation ................................................ 32
  3.8 The metric-covariant derivative ................................ 32
  3.9 Outlook: connections on fibre bundles ......................... 33
  3.10 Exercises .......................................................... 35

4 Differential forms 38
  4.1 Exterior algebra .................................................... 38
  4.2 The exterior derivative ............................................ 39
  4.3 The pull-back of differential forms ............................... 42
  4.4 Interior product .................................................... 43
  4.5 Lie derivative of differential forms .............................. 43
  4.6 The Cartan structural equations .................................. 45
  4.7 Exercises .......................................................... 48
5 Integration of differential forms
  5.1 The alternating symbol and the determinant . . . . . . . . . . 49
  5.2 The definition of the integral . . . . . . . . . . . . . . . . . 50
  5.3 Stokes’ Theorem . . . . . . . . . . . . . . . . . . . . . . . . . . 52
  5.4 The volume form on a Riemannian manifold . . . . . . . . . . 53
  5.5 The *-operator and the Hodge dual of a p-form . . . . . . . . . 54
  5.6 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

6 de Rham cohomology
  6.1 Closed and exact forms . . . . . . . . . . . . . . . . . . . . . . 56
  6.2 Integrals, the boundary operator $\partial$ and homology . . . . 59
  6.3 Homology and de Rham’s Theorem . . . . . . . . . . . . . . . . 61
  6.4 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

7 Symplectic Geometry
  7.1 Symplectic forms . . . . . . . . . . . . . . . . . . . . . . . . . 63
  7.2 Hamiltonian Systems . . . . . . . . . . . . . . . . . . . . . . . 66
  7.3 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . 69
1 Manifolds and tensors

1.1 Topological definitions

A topology on a given set \( S \) is a collection of open subsets, which are defined to be subsets with the following properties: 1. Any union of open sets is an open set. 2. Any finite intersection of open sets is an open set. 3. \( S \) and the empty set are open sets. A set with a topology is called a topological space. Its elements are usually called points.

A subset is called closed if its complement is open. An open neighbourhood of a point in \( S \) is an open subset of \( S \) that contains the point.

An open cover \( \{ O_\alpha \} \) (with \( \alpha \) in some index set) of \( S \) is a collection of open subsets whose union contains \( S \). A locally finite open cover of \( S \) is an open cover where every point in \( S \) has an open neighbourhood that intersects with only finitely many elements of the cover. A refinement of an open cover is a new open cover where each element of the new cover is contained in one of the old.

A topological space is called connected if the only subsets that are both open and closed are \( S \) and the empty set.

A topological space is called Hausdorff if any two points have non-intersecting open neighbourhoods.

A subset of a topological space is called compact if every open cover has a finite subcover.

A topological space is called paracompact if every open cover admits a locally finite refinement.

1.2 Manifolds

An \( n \)-dimensional manifold \( M \) is a Hausdorff topological space that can locally be identified with \( \mathbb{R}^n \). That means it can be covered by open neighbourhoods which map bijectively into open neighbourhoods of \( \mathbb{R}^n \):

\[
\varphi : M \supset U \rightarrow V \subset \mathbb{R}^n, \\
p \mapsto x^i := (x^1, x^2, \ldots, x^n).
\]

(1)

Such a map is called a chart, or a coordinate system. A collection of charts that covers the manifold is called an atlas. This definition becomes nontrivial where two charts overlap. We then have a map from \( \mathbb{R}^n \) to \( \mathbb{R}^n \)

\[
\varphi_1 : p \mapsto x_1^1, \\
\varphi_2 : p \mapsto x_2^1, \\
\psi = \varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2) \\
x_1^i \mapsto x_2^i.
\]

(2)

If all such coordinate transformations \( \psi \) are \( C^k \), we say the manifold is \( C^k \). The collection of all \( C^k \)-related charts is called the maximal atlas.

We now restrict to smooth (\( C^\infty \)) manifolds. If each \( \varphi \) in the atlas maps to a neighbourhood in \( \mathbb{R}^n \) with the same \( n \), we say the manifold has dimension \( n \).
1.3 Partitions of unity

An important technical tool for proving that local statements about manifolds hold globally is a **partition of unity**. This is a locally finite open cover \( \{O_\alpha\} \) together with a set of smooth (\( C^\infty \)) functions \( \{\phi_\alpha\} \) on the manifold such that

\[
\begin{align*}
\phi_\alpha(p) &= 0 \quad \text{for} \quad p \notin O_\alpha, \\
0 &\leq \phi_\alpha \leq 1, \\
\sum_\alpha \phi_\alpha &= 1.
\end{align*}
\]

(3) \hspace{1cm} (4) \hspace{1cm} (5)

Note that because the cover is locally finite, the sum is finite at any point.

**Theorem**: Every paracompact connected \( C^\infty \) Hausdorff manifold admits a partition of unity.

**Idea of proof**: We need paracompactness to get the locally finite open cover. From the smooth function on \( \mathbb{R} \)

\[
\phi(x) = \begin{cases}
0, & x < 0 \\
e^{-\frac{1}{x^2}}, & x > 0
\end{cases}
\]

(6)

we can construct a smooth function with support on an interval, hence a smooth function with support on a ball in \( \mathbb{R}^n \), and hence, using local coordinates, a smooth function on \( M \) with support in an open neighbourhood. Finally, we normalise to get the desired \( \phi_\alpha \).

1.4 Vectors

In going from \( \mathbb{R}^n \) to an \( n \)-dimensional manifold \( M \), we have kept the idea of coordinates (locally, each point \( p \) has unique coordinates \( x^i \)), but we lose the concept of subtracting two position vectors to get a direction vector: “\( p - q \)” is not defined, and \( x^i - y^i \) depends on the chart. Instead, we define vectors as derivative operators acting tangent to a curve.

Let

\[
c : I \rightarrow M \\
t \mapsto p(t)
\]

be a curve in \( M \) with parameter \( t \). In coordinates this is

\[
\varphi \circ c : I \rightarrow \mathbb{R}^n \\
t \mapsto x^i(t).
\]

(8)

Also, let

\[
\phi : M \rightarrow \mathbb{R} \\
p \mapsto \phi(p)
\]

(9)

be a function on \( M \).

We define the **tangent vector** \( X \) to the curve \( c \) at the point \( p = c(a) \) as a map from any function \( f \) to a number \( Xf \),

\[
X : f \mapsto Xf := \left. \frac{df}{dt} \right|_{t=a} f(c(t)).
\]

(10)
This definition of a vector as a derivative operator inherits the **Leibniz rule**
for the product of two functions from the usual derivative, namely

\[ X(fg) = fXg + gXf. \] (11)

Effectively, we have defined a vector as an equivalence class of curves going
through the point \( p \) “with the same velocity”. Therefore we need to show
explicitly that all such vectors at a point form a **vector space**. That is, we
need to show addition and multiplication by numbers.

**Idea of proof**: Say \( c_1 \) and \( c_2 \) are curves through \( p \) with tangent vectors \( X_1 \) and \( X_2 \). Then

\[
\tilde{c}_3 = \varphi \circ c_1 + \varphi \circ c_2 - \varphi(p)
\] (12)
is a curve in \( \mathbb{R}^n \) and

\[
c_3 = \varphi^{-1} \circ \tilde{c}_3
\] (13)
is a curve in \( M \) which goes through \( p \) with tangent vector \( X_1 + X_2 \).

In local coordinates

\[
\varphi \circ c : t \mapsto x^i(c(t)) =: x^i(t)
\] (14)
and

\[
\phi \circ \varphi^{-1} : x^i \mapsto \phi(p(x^i)) =: \phi(x^i).
\] (15)

Using the chain rule for a function of \( n \) variables,

\[
X\phi = \frac{d}{dt} \phi(x^i(t)) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x^i} \frac{dx^i}{dt} = \sum_{i=1}^{n} X_i \frac{\partial \phi}{\partial x^i},
\] (16)

where the \( n \)-tuple of numbers

\[
x^i := \frac{d}{dt} \bigg|_{t=a} x^i(c(t))
\] (17)
are the **coordinate components** of \( X \) in the chart \( \varphi \).

Under a change of coordinates

\[
X\phi = \sum_{i=1}^{n} X^i \frac{\partial \phi}{\partial x^i}
\]
\[
= \sum_{i=1}^{n} \left( \sum_{j=1}^{n} \frac{\partial x^i}{\partial x^j} \frac{dx^j}{dt} \right) \left( \sum_{k=1}^{n} \frac{\partial \phi}{\partial x^k} \frac{dx^k}{dt} \right)
\]
\[
= \sum_{j=1}^{n} \sum_{k=1}^{n} \left( \sum_{i=1}^{n} \frac{\partial x^i}{\partial x^j} \frac{\partial x^j}{\partial x^k} \right) \frac{dx^j}{dt} \frac{dx^i}{dt} \frac{\partial \phi}{\partial x^k}
\]
\[
= \sum_{j=1}^{n} \sum_{k=1}^{n} \delta^k_j \frac{dx^j}{dt} \frac{\partial \phi}{\partial x^k}
\]
\[
= \sum_{j=1}^{n} \frac{dx^j}{dt} \frac{\partial \phi}{\partial x^j},
\] (18)
and so in the new chart \( \varphi' \) the new components of \( X \) are

\[
X'^i = \frac{dx'^i}{dt} = \sum_{j=1}^{n} \frac{\partial x'^i}{\partial x^j} \frac{dx^j}{dt} = \sum_{j=1}^{n} \frac{\partial x'^i}{\partial x^j} X^j.
\]  

(19)

(The Kronecker delta \( \delta^i_j \) evaluates to 1 for \( i = j \) and 0 otherwise.)

From now on, we will use the **Einstein summation convention**, by which any pair of one index up and one index down is implicitly summed over without the need to write the sum explicitly. For example, the vector transformation law can be written as

\[
X'^i = \frac{\partial x'^i}{\partial x^j} X^j.
\]  

(20)

We can write the definition of a vector as a derivative operator more concisely as

\[
X = X^i \frac{\partial}{\partial x^i}
\]  

by removing the function \( f \), where once again summation is implied. We can then think of the \( \partial / \partial x^i \) for \( i = 1 \ldots n \) either as derivative operators, or simply as the **coordinate basis of vector fields** at a point.

A coordinate \( x^i \) of a fixed chart on a manifold is also a function. Expanding \( X \) in the corresponding coordinate basis, we have

\[
X x^i = X^j \frac{\partial}{\partial x^j} x^i = X^j \delta^i_j = X^i.
\]  

(22)

Sometimes, numbers and number-valued functions on \( M \) are called scalars and scalar fields to distinguish them from vectors. In these notes we use the word **scalar** more narrowly to denote a number that does not change under a change of basis. Functions on \( M \) that are not scalars in this sense are the coordinates \( x^i \), or vector components \( X^i \). Similarly, for us a vector is not any \( n \)-tuple of numbers, but one that transforms in a specific way under a change of basis.

1.5 Covectors

We define the **tangent space** \( T_pM \) to the manifold \( M \) at the point \( p \) as the vector space of all possible vectors at \( p \). Clearly it is \( n \)-dimensional. Define a **covector** \( \omega \) at \( p \) as a linear map on \( T_pM \):

\[
\omega : T_pM \rightarrow \mathbb{R}
\]

\[
X \mapsto \omega(X)
\]  

(23)

The space of all covectors at \( p \) is again a vector space, and is by definition the **dual space** of \( T_pM \). It is called \( T^*_pM \), the **cotangent space** of \( M \) at \( p \), and is again \( n \)-dimensional.

Equally, we can consider vectors as linear maps on covectors, and so you will also find the notations \( X(\omega) \), \( \omega \cdot X \) and \( \langle \omega, X \rangle \) in the literature to express this idea.

Linearity means that \( \omega(\alpha X + Y) = \alpha \omega(X) + Y \), and so if \( X^i \) are the components of \( X \) in the coordinate basis of a given chart \( \varphi \) there exist numbers \( \omega_i \) such that

\[
\omega(X) = \omega_i X^i.
\]  

(24)
Every function \( \phi \) on a manifold gives us a covector field \( d\phi \) as
\[
d\phi : T_pM \rightarrow \mathbb{R}
\]
\[
X \mapsto X\phi,
\]
(25)
called the gradient of \( \phi \). (You may have learnt to think about the gradient of a function as a vector, but it is a covector. We need a metric to define a gradient vector – see later.) In components
\[
d\phi(X) = X^i d\phi_i \quad \text{with} \quad (d\phi)_i = \frac{\partial \phi}{\partial x^i}.
\]
(26)

In particular, locally we can choose \( \phi \) to be the \( n \) coordinate functions \( x^i \) of a chart, and we have
\[
dx^i : T_pM \rightarrow \mathbb{R}
\]
\[
X \mapsto X x^i = X^i
\]
(27)
because of (22). The coordinate basis \( dx^i \) of covectors is dual to the coordinate basis \( \partial/\partial x^i \) of vectors in the sense that
\[
dx^i \left( \frac{\partial}{\partial x^j} \right) = \frac{\partial x^i}{\partial x^j} = \delta^i_j.
\]
(28)
Using this basis we can expand any covector \( \omega \) as \( \omega = \omega_i dx^i \) and we have
\[
\omega(X) = (\omega_i dx^i) \left( X^j \frac{\partial}{\partial x^i} \right) = \omega_i X^j dx^i \left( \frac{\partial}{\partial x^j} \right) = \omega_i X^j \delta^i_j = \omega_i X^i,
\]
(29)
so the \( \omega_i \) we defined in (24) are the components of \( \omega \) in the coordinate basis \{\( x^i \}\).

Note that index position always matters in differential geometry: coordinate indices and vector component indices are up, covector component indices are down. \( \partial/\partial x^i \) also has its index down. Only index pairs where one index is up and one is down are ever contracted under the Einstein summation convention.

Note also that the \( i \)-th coordinate component of the vector \( \partial/\partial x^j \) is \((\partial/\partial x^j)^i = \delta^i_j\). (The vector index is up, but the index labelling the vector as a coordinate basis vector is down.) Conversely, the \( i \)-th coordinate component of the covector \( dx^j \) is \((dx^j)_i = \delta^j_i\).

Under a change of coordinates
\[
\omega(X) = \omega_i X^i = \omega_i \left( \frac{\partial x^j}{\partial x'^i} X'^j \right) = \left( \omega_i \frac{\partial x^j}{\partial x'^i} \right) X'^j =: \omega'_j X'^j,
\]
(30)
so covectors transform as
\[
\omega'_j = \frac{\partial x^j}{\partial x'^i} \omega_i.
\]
(31)

1.6 Tensors

Now we define a tensor of rank \((r, s)\) as a multilinear map on \( r \) covectors and \( s \) vectors,
\[
T : (T^*_pM)^r \times (T_pM)^s \rightarrow \mathbb{R}
\]
\[
\omega_1, \omega_2, \ldots, \omega_r, X_1, X_2, \ldots X_s \mapsto T(\omega_1, \omega_2, \ldots, \omega_r, X_1, X_2, \ldots X_s)
\]
(32)
where
\[ T(\ldots, \alpha \omega + \mu, \ldots) = \alpha T(\ldots, \omega, \ldots) + T(\ldots, \mu, \ldots). \] (33)
and so for every “slot” of the tensor \( T \). Note that the order of the slots matter.

From multilinearity it follows that there are numbers, called **tensor components** \( T_{ij...jk...l} \) such that, for example the example of a \((2, 1)\)-tensor \( T \),
\[ T(\omega, \mu, X) = T^{ij}_{\ k} \omega_i \mu_j X^k, \] (34)
and similarly for tensors of other rank. (See later for an explanation of why I prefer to order all tensor indices irrespective of their being up or down, i.e. why I write \( T^{ij}_{\ k} \) rather than \( T^{ij}_{\ k} \).)

Clearly tensor components transform as
\[ T'_{ij...k...} = \frac{\partial x'_{i}}{\partial x^p} \frac{\partial x^q}{\partial x_{j}} \cdots \frac{\partial x^n}{\partial x_{k}} \cdots T_{lm...}^{...}, \] (35)
Here the dots represent an additional factor on the right-hand side for each additional index on \( T \): of the form \( \partial x'_{\ i} / \partial x^i \) for up indices, and of the form \( \partial x_{\ i} / \partial x'_{i} \) for down indices: it should be clear how all the indices match up. In particular, a vector is a \((1, 0)\) tensor, a covector is a \((0, 1)\) tensor, and a scalar (in the strict sense of the word) is a \((0, 0)\) tensor. Clearly, all tensors of the same rank form a vector space.

From the transformation properties of tensors, it follows that a \((1, 1)\)-tensor can also be seen as a map \( T_{p}^{\ k} M \rightarrow T_{p}^{\ k} M \) or \( T^{*}_{\ p} M \rightarrow T^{*}_{\ p} M \) (see exercise). Tensors of other types have similar interpretations, for example a \((1, 3)\)-tensor can be interpreted as a map from 3 vectors into 1 vector, from 2 vectors into a \((1, 1)\)-tensor, and so on.

1.7 Contraction of tensors
\( \omega(X) = \omega_{i} X^i \) is the simplest example of the **contraction** of two tensors. The result in this case is a scalar. But we can always contract an up index on one tensor with a down index on another tensor to obtain a new tensor. The same thing can be done with more than one pair of indices. To save notation, we present an example rather than the general case. Take two tensors \( S \) and \( T \) whose components in a given chart are \( S_{ij} \) and \( T_{kl} \). Define a new \((1, 1)\) tensor \( S \cdot T \) by defining its components as
\[ (S \cdot T)^{ij}_{\ k} = S_{ik} T^{kj}, \] (36)
again using the summation convention.

We still need to show that this really is a tensor, meaning that it transforms correctly. Transforming each tensor separately, we have
\[ S_{ik} T^{kj} = \left( S'_{pq} \frac{\partial x'^p}{\partial x^i} \frac{\partial x'^q}{\partial x^k} \right) \left( T'^{rs}_{\ tr} \frac{\partial x^r}{\partial x'^s} \frac{\partial x^t}{\partial x'^r} \right). \] (37)
From the chain rule for partial derivatives we have
\[ \frac{\partial x'^q}{\partial x^k} \frac{\partial x^k}{\partial x'^r} = \delta^q_{r}, \] (38)
and so
\[ S_{ik} T^{kj} = \frac{\partial x'^p}{\partial x^i} \frac{\partial x^j}{\partial x'^r} \left( S'_{pr} T'^{rs} \right), \] (39)
which is the correct transformation for a \((1, 1)\) tensor.
1.8 Non-coordinate bases

There is no need to restrict ourselves to a coordinate basis of vectors and covectors. We can consider a general basis $\{e_a\}$, $a = 1 \ldots n$ of the tangent space. A non-coordinate basis is also called a moving frame in the literature. For any given basis $\{e_a\}$ of vector fields, the dual basis $\{\theta^b\}$ of covector fields is defined uniquely by

$$\theta^b(e_a) = \delta^b_a.$$  

(The dual basis of the coordinate basis $\partial/\partial x^i$ of vectors is of course the coordinate basis $dx^i$ of covectors.)

Under a change of basis of the tangent space

$$e_a \rightarrow e'_a = A^a_b e_b,$$

where $A^a_b$ is an invertible matrix, the dual basis of the cotangent space must therefore transform as

$$\theta^a \rightarrow \theta'^a = (A^{-1})^a_b \theta^b,$$

and tensor components in these bases transform in the obvious way. Note the order and position of indices on $A$ and $A^{-1}$. We see that the Jacobi matrix $\partial x'^i/\partial x^j$ is just a particular instance of a basis transformation matrix $A^a_b$.

1.9 Abstract index notation

$\omega \cdot X$ can only mean $\omega^a X_a$, but there are four ways of contracting the $(0,2)$ and $(2,0)$ tensors $S$ and $T$ to obtain a $(1,1)$ tensor, so $S_{ac} T^{cb}$ is clearer notation than $S \cdot T$ because it tells us which index pair has been contracted. Often, the best notation to use in tensor equations is $S_{ac} T^{cb}$, but considering this not as a collection of tensor components in a specific basis but rather as a notation for the tensor as a whole. This is called the abstract index notation. Note that if we write a tensor equation (such as (36)) in abstract index notation, then all dummy indices should come in pairs, with no pair of dummy indices repeated to avoid ambiguity, while the free indices must be the same ones, in the same position, on the left-hand and right-hand side. However, it is the same tensor equation if we rename any free index or dummy index pair.

As an example of abstract index notation, $T_{ab} S_b$ is the same tensor as $T_{ba} S_a$ (relabelling of free and dummy indices) but is not the same tensor as $T_{ba} S_b$ (a different contraction).

We can define the antisymmetric part of a tensor with respect to a group of two or more vector or covector indices. Giving two examples for simplicity, we define

$$X_{[ab]} := \frac{1}{2!} (X_{ab} - X_{ba}),$$

$$X_{[abc]} := \frac{1}{3!} \left( X_{abc} + X_{bca} + X_{cab} - X_{bac} - X_{acb} - X_{cba} \right),$$

and so on. Generally $p!$ terms are needed to antisymmetrise over $p$ indices. Similarly, round brackets around a group of tensor indices denote symmetric part (using only plus signs). These would be more cumbersome to write without abstract index notation. We note for future use that a totally antisymmetric
tensor with all its \( k \) indices down is called a \( k \)-form, so a covector is a 1-form, \( \omega_{ab} = -\omega_{ba} \) is a 2-form, and so on.

In these lecture notes, we use \( abc \ldots \) for free and dummy indices that can be interpreted as either abstract indices or non-coordinate basis indices, and we use \( ijk \ldots \) for coordinate indices. We do this to stress that some equations are valid only in a coordinate basis, while others are true tensor equations. However, this is not a universal convention. Most books either consider only coordinate bases, or leave the distinction between abstract and coordinate indices to the context.

1.10 Tensor fields, and the commutator of vector fields

If we put a tensor (for example a vector) at each point of the manifold, and its coordinate components are \( C^k \) functions of the coordinates, then we have a \( C^k \) tensor field. We can now consider derivatives of tensor fields.

To make our formulae more compact, from now on we usually write

\[
\frac{\partial \phi}{\partial x^i} =: \phi_i
\]

for the partial derivative of a function with respect to a given coordinate system. Note that partial derivatives commute, or

\[
\phi_{ij} = \phi_{ji}
\]

(for manifolds of dimension \( \geq 2 \)).

How do we know a basis of vector fields is a coordinate basis? For any two vector fields \( X \) and \( Y \) we define a new vector field called their commutator or Lie bracket by its action on functions \( \phi \). \( X \phi \) and \( Y \phi \) are again functions, and we define

\[
[X,Y]\phi := X(Y\phi) - Y(X\phi)
\]

for all scalar fields \( \phi \). In coordinates

\[
[X,Y]\phi = X^i(Y^j\phi_j)_i - Y^i(X^j\phi_j)_i
\]

\[
= X^i(Y^j\phi_{ji} + Y^j\phi_{ij}) - Y^i(X^j\phi_{ji} + X^j\phi_{ij})
\]

\[
= (X^iY^j - Y^iX^j)\phi_{ij},
\]

and so

\[
[X,Y]^i = X^iY^j - Y^jX^i = XY^i - YX^i,
\]

As \( [X,Y] \) is by definition a vector field, this transform as a vector under a change of coordinates (see exercise).

We define the structure constants of a non-coordinate basis by

\[
\{e_b,e_c\} =: C^{ac}_{bc} e_a.
\]

The structure constants of a coordinate basis vanish identically.

We can then write the components of the Lie bracket in a non-coordinate basis:

\[
[X^ae_a, Y^be_b]\phi = (X^ae_aY^b - Y^ae_aX^b)e_b\phi + X^aY^b\{e_a,e_b\}\phi
\]

and so

\[
[X,Y]^c = (XY^c - YX^c) + X^aY^bC^{ac}_{ab}.
\]
(Note that the individual terms on the right-hand side do not transform as components of a tensor, so while this defines a tensor, it is not a tensor equation.) In $XY^c$, the vector $X$ acts as a derivative operator on the function $Y^c$ (the component of $Y$ in a fixed basis).

From (49), or directly from the commutation of partial derivatives, we see that in a chart $\varphi$

$$\left[ \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] = 0 \quad (53)$$

Conversely, if we have $n$ vector fields on a neighbourhood of a point that all commute we can locally find coordinates such that these vector field are the partial derivatives with respect to these coordinates. (The vanishing of the commutators are precisely the integrability conditions required for the construction of the $x^i$ by integration along any curve.)

The Lie bracket obeys the Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0. \quad (54)$$

1.11 The metric tensor

A key tensor field is the metric tensor. Recall that 3-dimensional Euclidean space comes equipped with an inner product or dot product of vectors, which in Cartesian coordinates is given by

$$X \cdot Y := X^1Y^1 + X^2Y^2 + X^3Y^3 \quad (55)$$

(In this equation only, the dot is used as the symbol for the inner product.)

A manifold does not a priori come with such an inner product, but we can introduce the inner product of any two tangent vectors at any point by defining

$$g(X, Y) := X^ag^b_{ab} \quad (56)$$

As $X$ and $Y$ are vectors, the metric tensor $g_{ij}$ must be a (0,2)-tensor in order for the inner product to be a scalar (invariant under basis changes). As we require $g(X, Y) = g(Y, X)$, we must have

$$g_{ab} = g_{ba} \quad (57)$$

that is, the metric tensor is symmetric.

Like the Euclidean inner product, the metric gives us the length or absolute value of a vector

$$|X| := \sqrt{g(X, X)} \quad (58)$$

and the angle $\theta$ between two (non-zero) vectors by

$$\cos \theta := \frac{g(X, Y)}{|X||Y|}. \quad (59)$$

An equivalence class of metrics up to a position-dependent factor is called a conformal metric. It defines angles but not lengths.

The inner product on Euclidean space $\mathbb{R}^3$ is given by the Euclidean metric $\gamma_{ij}$. Euclidean space has a system of global coordinates $x^i$ called the Cartesian coordinates in which $\gamma_{ij}$ has components $\delta_{ij}$. Its components in any other
coordinate system (say spherical or cylindrical coordinates) can be computed from the tensor transformation law.

More generally, any symmetric tensor (0,2)-tensor that is invertible can be considered a metric. We define the inverse metric $g^{ab}$ by defining

$$g^{ab} g_{bc} := \delta^a_c,$$  \hspace{1cm} (60)

where the (1,1)-tensor $\delta^a_b$ represents the identity operator, which happens to have the same components in any basis. This is to say that in any basis, $g^{ab}$ written out as a (symmetric) matrix is the matrix inverse of $g_{ab}$. (See exercise).

Note that if we treat the metric tensor in any given coordinate system as a matrix, then it is positive definite if all its eigenvalues are positive, and invertible if all its eigenvalues are non-zero, and this property is independent of the coordinate system. (Idea of proof: the coordinate transformation matrix is itself invertible, and so cannot have zero eigenvalues).

The Euclidean metric, for example, has the additional property that it is positive definite, meaning that \[ X^a X^b g_{ab} > 0 \] for any $X^a \neq 0$. \hspace{1cm} (61)

Any positive definite metric is called a Riemannian metric, and a manifold equipped with a Riemannian metric is called a Riemannian manifold. A metric with both positive and negative (but no zero) eigenvalues is called pseudo-Riemannian. The main example of a pseudo-Riemannian metric is the metric on spacetime in relativistic physics, which has three positive (space) and one negative (time) eigenvalues.

We can write any (0,2)-tensor in terms of a basis made from the tensor product of basis covectors, or

$$g = g_{ab} \theta^a \otimes \theta^b,$$  \hspace{1cm} (62)

where the summation convention applies. In a coordinate basis

$$g = g_{ij} dx^i \otimes dx^j.$$  \hspace{1cm} (63)

This acts on a pair of vectors as

$$g(X,Y) = g_{ij} (dx^i \otimes dx^j)(X,Y) = g_{ij} dx^i(X) dx^j(Y) = g_{ij} X^i X^j.$$  \hspace{1cm} (64)

For the metric tensor (but no other tensor), (63) is often written as

$$ds^2 = g_{ij} dx^i dx^j.$$  \hspace{1cm} (65)

This notation is called the line element because we can define the length of the curve $c(t)$ with tangent vector $X(t)$ between $t = a$ and $t = b$ as

$$s := \int_a^b \frac{ds}{dt} dt := \int_a^b \sqrt{g(X(t), X(t))} dt = \int_a^b \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt.$$  \hspace{1cm} (66)

On a Riemannian or pseudo-Riemannian manifold, one can use the metric to raise or lower indices on a given tensor. For example, we can define the gradient vector

$$\nabla^a f := g^{ab} (df)_b = g^{ab} e_b f = g^{ab}(e_b)^i f_i.$$  \hspace{1cm} (67)
or just
\[ \nabla^i f := g^{ij} f_j \quad (68) \]
in a coordinate basis.

The raising and lowering of indices is often carried out \textbf{implicitly}, that is, if a vector \( T^a \) is defined, one defines the covector \( T_a \) as
\[ T_a := g_{ab} T^b, \quad (69) \]
and similarly for any indices on any tensor, always using the same base symbol (here \( T \)) for what is then considered to be “the same tensor with indices moved”.

First warning: If we use the convention of moving tensor indices implicitly, the order of indices matters irrespective of position. Therefore, to avoid ambiguity, it is safest to always write \( T_{abc} \) (\textit{un consolidated form}), rather than \( T^{abc} \) (\textit{consolidated form}), in case we implicitly define, for example, \( T_{abc} \). (The Riemann tensor will be an example of this later on.)

Second warning: On Euclidean space in a Cartesian coordinate basis, the components of the metric \( \gamma \) are \( \gamma_{ij} = \delta_{ij} \) and its inverse \( \gamma^{ij} = \delta^{ij} \), and so one has, in this coordinate basis only, \( T^i = T_i \). Physics or engineering textbooks or papers often use this to write all tensor indices on \textbf{Cartesian tensors} as down indices. This is potentially confusing.

If we define
\[ g^{ab} = \delta^a_b, \quad (70) \]
the implicit raising and lowering of indices applied to the metric tensor itself is consistent with (60). A consequence of (70) is
\[ g_a^a = n, \quad (71) \]
so that the trace of the metric tensor is the dimension of the manifold.

On a Riemannian manifold, it may be useful to work in an \textbf{orthonormal basis} of vector fields defined by
\[ g_{ab} = g(e_a, e_b) = \delta_{ab}. \quad (72) \]
Such a basis cannot in general be a coordinate basis. As an example, consider Euclidean space \( E^2 \) in Cartesian and polar coordinates. The line element is
\[ ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2 \quad (73) \]
The Cartesian coordinate basis
\[ e_x = \frac{\partial}{\partial x}, \quad e_y = \frac{\partial}{\partial y} \quad (74) \]
is already orthonormal. The polar coordinate basis is already orthogonal (because \( g_{r\theta} = 0 \)) but not normalised. An orthonormal basis is
\[ e_r = \frac{\partial}{\partial r}, \quad e_\theta = \frac{1}{r} \frac{\partial}{\partial \theta}. \quad (75) \]
(We use round brackets here to indicate specific non-coordinate indices. This is a useful but not a standard convention.)
One consequence of the existence of a partition of unity is that the manifold admits a Riemannian metric:

**Theorem:** Every paracompact manifold admits a Riemannian metric.

**Proof:** From paracompactness we have the existence of a locally finite cover \( \{ O_\alpha \} \) and of a partition of unity \( \{ \phi_\alpha \} \). To construct a Riemannian metric explicitly, we note that in local coordinates \( g_{ij} = \delta_{ij} \) is a positive definite metric, and we define local metrics \( g_\alpha \) in this way. We can then define the global \( C^\infty \) Riemannian metric

\[
g = \sum_\alpha \phi_\alpha \, g_\alpha,
\]

where we have used the obvious fact that the sum of two positive definite metrics is positive definite.

We can use the same method to construct smooth scalar and tensor fields on a manifold.

### 1.12 Exercises

1. Explain carefully how our definition of a manifold fixes its dimension.

2. How could one define a scalar function on \( M \) to be \( C^k \)?

3. Complete the proof that tangent vectors to curves form a vector space (addition, multiplication by scalars).

4. Give two local coordinate systems on the 2-sphere \( S^2 \). Give an example of one vector field in both coordinate systems.

5. Prove directly that (49) actually transforms as the components of a vector field.

6. Prove the Jacobi identity (54).

7. Show that if the components of a tensor are \( \delta^i_j \) (meaning 1 if \( i = j \), and 0 otherwise) in one coordinate system, they are \( \delta^j_i \) in every coordinate system.

8. Explain in what sense a \((1,1)\)-tensor \( M^{ab} \) can be considered as a map \( TM \to TM \), or equivalently as a matrix. What are the components of the unit tensor?

9. Write the line element of Euclidean space \( \mathbb{R}^3 \) in spherical polar coordinates, and find orthogonal bases of vectors and covectors (dual to each other) aligned with the coordinate bases.

10. Show that the gradient vector \( \nabla^a f \) gives the direction of steepest descent (the direction in which the rate of change of \( f \) with distance is largest).
2 Maps of manifolds, integral curves and the Lie derivative

2.1 Motivation: derivatives of vector fields

We would like to define derivatives of tensor fields, for example of a vector field \( X^i \). One could try the directional derivative of \( X^i \) along the vector \( Y \),

\[
Y^j \frac{\partial X^i}{\partial x^j},
\]

but this object does not transform as a vector field. Or one could try

\[
\frac{\partial X^i}{\partial x^j},
\]

but this does not transform as a \((1,1)\) tensor.

However, genuine tensors with similar properties can be defined, and they are the Lie derivative, the covariant derivative, and the exterior derivative, respectively. All have geometric significance, and hence applications in physics. All have some restrictions: the Lie derivative along \( Y \) requires knowledge of \( Y^j, i \), so it is not quite a derivative at one point. The covariant derivative requires an additional geometric structure on the manifold, called a connection. Finally, the exterior derivative requires neither, but only works on totally antisymmetric tensors with all indices down, also called differential forms. We consider these in the next sections, starting here with the Lie derivative.

Behind the Lie derivative is the idea of a map from \( M \) to \( M \) and which depends differentiably on a parameter \( t \), so that it “moves points along”. We use this to define the Lie derivative along the tangent vector to a curve. However, for a clear presentation we start with maps from \( M \) to another manifold \( N \), and only later set \( N = M \).

2.2 Maps of manifolds

Let \( M \) and \( N \) be two manifolds, with dimensions \( m \) and \( n \) and

\[
f : M \to N
\]

\[
p \mapsto f(p)
\]

a map from \( M \) to \( N \). In local charts on each manifold,

\[
\varphi : M \to \mathbb{R}^m
\]

\[
p \mapsto x^i
\]

and

\[
\varphi' : N \to \mathbb{R}^n
\]

\[
q \mapsto y^\alpha
\]

we have

\[
\varphi' \circ f \circ \varphi^{-1} : x^i \mapsto y^\alpha.
\]
The map $f$ is called $C^k$ if $\varphi' \circ f \circ \varphi^{-1}$ is, and its rank is defined to be the rank of $\varphi' \circ f \circ \varphi^{-1}$.

$f$ is called an immersion if rank $f = m$. $f$ is called an embedding if it is an immersion, is injective and is a homeomorphism of $M$ onto $f(M)$, with the topology of $f(M)$ induced by the topology of $N$. $f$ is called a submersion if rank $f = n$.

For a function (scalar field) $\phi$ on $N$

$$\phi : N \to \mathbb{R}$$

$$q \mapsto \phi(q)$$

we define its pull-back to $M$ by

$$f^* \phi : M \to \mathbb{R}$$

$$p \mapsto (f^* \phi)(p) := \phi(f(p)).$$

In local coordinates this means, in relaxed notation, that

$$(f^* \phi)(y^\alpha) = \phi(x^i(y^\alpha)).$$

The map $f$ is defined to be $C^\infty$ if the pull-back of every $C^\infty$ function $\phi$ is also $C^\infty$.

For a vector field $V$ on $M$ we can then define its push-forward to a vector field on $N$ by giving its action on any function $\phi$ on $N$, in the obvious way:

$$(f_* V) \phi := V(f^* \phi)$$

In local coordinates this equation is

$$(f_* V)^{\alpha} \frac{\partial \phi}{\partial y^\alpha} = V^i \frac{\partial}{\partial x^i} \phi(y(x)) = V^i \frac{\partial y^\alpha}{\partial x^i} \frac{\partial \phi}{\partial y^\alpha},$$

and as this holds for any $\phi$, we must have

$$(f_* V)^{\alpha} = V^i \frac{\partial y^\alpha}{\partial x^i}.$$ (88)

(That a vector field can be pushed forward is also intuitively clear if we think of it as the tangent vector to a curve that is pushed forward point by point by $f$.)

Next we can define the pull-back of a covector from $N$ to $M$ through

$$(f^* \omega)V := \omega(f_* V).$$

In local coordinates

$$(f^* \omega)_i V^i := \omega_\alpha \left( \frac{\partial y^\alpha}{\partial x^i} V^i \right),$$

and so

$$(f^* \omega)_i = \frac{\partial y^\alpha}{\partial x^i} \omega_\alpha.$$ (91)

(Again intuitively, covectors can be written as linear combinations of the gradients of functions (for example $dx'$), and we have already seen that functions are pulled back.)
Now if \( f \) is actually one-to-one (which implies that \( m = n \)) then \( \partial x^i / \partial y^\alpha \) also exists (it is the matrix inverse of \( \partial y^\alpha / \partial x^i \)) and then we can define both the push-forward and the pull-back of tensors of any type, for example the push-forward

\[
(f_* T)^\alpha_\beta = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} T^i_j, \tag{92}
\]

and similarly for other cases – it is always obvious what Jacobian matrices must be used.

### 2.3 Integral curves and 1-parameter families of maps

**Definition:** An **integral curve** of a vector field \( V \) is a curve in \( M \) such that its tangent vector at each point is \( V \).

**Theorem:** A smooth nonvanishing vector field locally defines a unique integral curve \( f_p(t) \) through any point \( p \), such that \( f_p(0) = p \).

**Proof:** Let \( x_i^p \) be the coordinates of \( p \), and let \( x_i^t \) be the coordinates of \( f_p(t) \). Then \( x_i^t \) obeys

\[
\frac{dx_i}{dt} = X^i(x^m(t)), \quad x_i^0 = x_i^p \tag{93}
\]

which is a system of \( n \) first-order ordinary differential equations. By a standard theorem, these have a unique solution locally.

Changing our point of view, the collection of all integral curves defines a one-to-one map from \( M \) to itself which depends on a parameter \( t \) and is at least once differentiable in \( t \), with \( f_0 = \text{Id} \).

**Definition:**

\[
f_t : M \times \mathbb{R} \to M \quad (p, t) \mapsto f_t(p) := f_p(t). \tag{94}
\]

In local coordinates, we shall write

\[
(x^i, t) \mapsto \tilde{x}^i(x, t). \tag{95}
\]

### 2.4 Lie derivative of a function

We can now define the Lie derivative of the function (scalar field) \( \phi \) at the point \( p \) along the vector field \( V \) as

\[
\mathcal{L}_V \phi := \lim_{t \to 0} \frac{1}{t} [f_t^* \phi - \phi]. \tag{96}
\]

\( f_t(p) \) for fixed \( p \) is an integral curve of \( V \). Hence

\[
(\mathcal{L}_V \phi)(p) = \lim_{t \to 0} \frac{1}{t} [\phi(f_t(p)) - \phi(p)] = \frac{d}{dt} \phi(f_t(p)) \bigg|_{t=0} = V_p \phi, \tag{97}
\]

and so \( \mathcal{L}_V \phi = V \phi \).
In preparation for the next calculation, let us work this out explicitly in coordinates. Expanding \( \tilde{x}^i(x, t) \) in \( t \), we have

\[
\tilde{x}^i(x, t) = x^i + V^i t + o(t)
\]

(98)

where the first equality holds because by assumption \( f_t \) is at least once differentiable in \( t \). Substituting the definition of \( (f_t)^* \phi \), we have

\[
\mathcal{L}_V \phi = \lim_{t \to 0} \frac{1}{t} \left[ \phi(\tilde{x}(x, t)) - \phi(x) \right]
\]

(99)

where in the third equality we have used the fact that \( \phi \) is at least once differentiable.

### 2.5 Lie derivative of a vector field

The basic problem with defining the derivative of a vector field is that we need to take the difference of that vector field at two different points, but that this difference is not defined on a manifold. However, we can pull back the vector at \( q = f_t(p) \) to a vector at \( p \) and so subtract two vectors at \( p \). Actually, this pull-back is the push-forward with \( f_t^{-1} \). This gives rise to the following

**Definition:**

\[
\mathcal{L}_V W := \lim_{t \to 0} \frac{1}{t} \left[ W(f_t(p)) - W(p) \right].
\]

(100)

**Theorem:**

\[
\mathcal{L}_V W = [V, W].
\]

(101)

We will give two proofs of this: the first one by direct calculation, the second one more elegant (later).

**Proof:** \( f_t^{-1} \) can be expanded in coordinates as (compare (98))

\[
x^i = \tilde{x}^i - V^i t + o(t) = \delta^i_j \tilde{x}^j - V^i t + o(t)
\]

(102)

and so the matrix representing \( (f_t^{-1})_* \) is

\[
\frac{\partial x^i}{\partial \tilde{x}^j} = \delta^i_j - \frac{\partial V^i}{\partial \tilde{x}^j} t + o(t) = \delta^i_j - \frac{\partial V^i}{\partial x^j} t + o(t).
\]

(103)

Substituting this last result, and the Taylor expansion of \( W(f_t(p)) \),

\[
(\mathcal{L}_V W)^i = \lim_{t \to 0} \frac{1}{t} \left[ \left( \delta^i_j - \frac{\partial V^i}{\partial \tilde{x}^j} t + o(t) \right) \left( W^j + \frac{\partial W^j}{\partial x^k} V^k t + o(t) \right) - W^i \right]
\]

\[
= \lim_{t \to 0} \frac{1}{t} \left[ W^i + \frac{\partial W^j}{\partial x^k} V^k t - \frac{\partial V^i}{\partial \tilde{x}^j} W^j t + o(t) - W^i \right]
\]

\[
= V^j W^i_j - W^j V^i_j,
\]

(104)
where we have relabelled a dummy index in the last equality. The first term is
the directional derivative (77) one would expect. The second term is unexpected,
but only both terms together transform as a vector.

2.6 Lie derivative of a covector field

To define the Lie derivative of a covector field, we define the Lie derivative to
obey the product rule (which is one of the axioms of any derivative operator)
for products of tensors:

\[ \mathcal{L}_V (T \cdot S) := \mathcal{L}_V T \cdot S + T \cdot \mathcal{L}_V T, \]  \hspace{1cm} (105)

where the dot stands for any contractions between the tensors. In particular,
this gives us

\[ \mathcal{L}_V (\omega \cdot W) = V^i (\omega_j W^j)_j = V^i \omega_j W^j + V^i \omega_j W^j, \]
\[ = \omega (\mathcal{L}_V W) + (\mathcal{L}_V \omega) \cdot W \]
\[ = \omega_j (V^i W^j,i - W^i V^j,j) + (\mathcal{L}_V \omega)_i W^i. \]  \hspace{1cm} (106)

As this must hold for any \( W^i \), we must have

\[ (\mathcal{L}_V \omega)_i = V^j \omega_{i,j} + V^j \omega_j. \]  \hspace{1cm} (107)

Again, the first term is expected, but only both terms together transform as a
covector.

2.7 Lie derivative of a tensor field of arbitrary rank

The product rule argument can be used on tensors of arbitrary rank, by con-
tracting each slot with a vector or covector. We find

\[ (\mathcal{L}_V T)_{ij \ldots} = V^k T_{ij \ldots,k} = V^k T_{ij \ldots} - V^i T_{jk \ldots} - \cdots + V^k T_{ij \ldots} + \cdots \]  \hspace{1cm} (108)

where the first term is the expected one, and the dots indicate that one correction
term is needed for each vector and covector index. As \( \mathcal{L}_V T \) is always a tensor
of the same type as \( T \), we will no longer write the brackets on the left-hand side
of (108).

Note that for the definition of the Lie derivative, \( V \) needs to be a \( C^1 \) vector
field, not just a vector at one point.

2.8 Coordinates adapted to a vector field

**Theorem:** If \( V \) is a smooth, non-vanishing vector field, there exists a coordinate
system \( x^i \) in the neighbourhood of a point \( p \) such that

\[ V = \frac{\partial}{\partial x^1}. \]  \hspace{1cm} (109)

**Idea of proof:** Choose a hypersurface \( S \) through \( p \) such that \( V \) is nowhere
tangent to it. Introduce coordinates \( x^2, x^3, \ldots \) on \( S \). Construct the integral
curve of \( V \) through each point on in a neighbourhood of \( p \) on \( S \). Label each
point on the integral curve by \( x^1 = t \) (the parameter of the integral curve) and
Such coordinates are useful for calculations and proofs. As an example, we prove once again the theorem that $\mathcal{L}_V W = [V, W]$. We work in coordinates such that (109) holds. In these coordinates the integral curves $f_t(p)$ are given by

$$\varphi \circ f_t : (x^1; t) = (x^1, x^2, \ldots, x^n; t) \mapsto \tilde{x}^i = (x^1 + t, x^2, \ldots, x^n)$$

(110)

Hence the matrix $\frac{\partial \tilde{x}}{\partial x}$ representing the action of $(f^{-1}_t)$ on tensor indices is the identity matrix for all $t$ and

$$\left(\mathcal{L}_V W\right)^i = \lim_{t \to 0} \frac{1}{t} \left[ W^i(x^1 + t, x^2, \ldots) - W^i(x^1, x^2, \ldots) \right] = \frac{\partial W^i}{\partial x^1} = V(W^i)$$

(111)

The last expression in this equation does not transform as a vector, so cannot be the correct expression in general. However, in the same special coordinates, $V^i = \delta^i_1$, and so

$$W(V^i) = 0$$

(112)

for any vector field $W$. Therefore, still, in these coordinates, we can write

$$\left(\mathcal{L}_V W\right)^i = V(W^i) - W(V^i) = [V, W]^i.$$  

(113)

But now both the left-hand side and the right-hand side define a vector, and so we have a tensor equation, which must hold in arbitrary coordinates.

More generally, if $V = \partial/\partial x^1$ in one particular coordinate system, then $V^i = \delta^i_1$ and so $\partial V^i/\partial x^j = 0$ in that coordinate system, and so, from (108), in these coordinates only,

$$\mathcal{L}_V T^{i\ldots j\ldots} = \frac{\partial}{\partial x^1} T^{i\ldots j\ldots}, \quad V = \frac{\partial}{\partial x^1}.$$  

(114)

In this sense, the Lie derivative is a generalisation of the partial derivative.

### 2.9 Submanifolds

**Definition:** An $n$-dimensional submanifold $N$ of an $m$-dimensional manifold $M$ is a subset of $M$ such that in a neighbourhood of any point $p$ in $N$ there exist local coordinates $x^1, \ldots, x^m$ such that $N$ is defined by $x^{n+1} = \cdots = x^m = 0$. The topology of $N$ is the topology induced by the topology of $M$. $m - n$ is called the codimension of $N$ in $M$. A submanifold of codimension 1 is called a hypersurface.

In fact, a null set of any $m - n$ functions, which do not have to be coordinates, defines a submanifold:

**Theorem:** A subset $N$ of $M$ defined by $n - m$ equations

$$\phi_{n+1}(p) = \cdots = \phi_m(p) = 0$$

(115)

where the $\phi_i$ are $C^1$ functions on $M$ and the map

$$M \to \mathbb{R}^{m-n}$$

$$p \mapsto (\phi_{n+1}(p), \ldots, \phi_m(p))$$

(116)
has rank $m - n$ is a differentiable submanifold of dimension $n$.

**Proof:** Let $y^i, i = 1, \ldots, n$ be any local coordinates. The Jacobian $\frac{\partial \phi^k}{\partial y^i}$ is an $(n - m) \times n$ matrix. If it has rank $n - m$, there must be a $(n - m) \times (n - m)$ submatrix that is invertible. We can re-order the coordinates $x^i$ so that this invertible square submatrix is $\frac{\partial \phi^k}{\partial y^i}$ for $i, k = n + 1, \ldots, m$. The inverse function theorem then tells us that we can locally invert $\phi^k(y^i)$ to get $y^i(\phi^k)$.

The coordinates in the definition are then

$$x^1 = y^1, \ldots, x^n = y^n, \quad x^{n+1} = \phi_{n+1}, \ldots, x^m = \phi_m. \quad (117)$$

and the map from $x^i$ to $y^i$ is invertible.

Clearly, the vector fields $\partial/\partial x^1, \ldots, \partial/\partial x^n$ span the tangent space of $N$ at $p$ and commute. Moreover, as the $x^i$ are coordinates on $M$ in a neighbourhood of $N$, the sets $x^i = c^i$ for $i = 1, \ldots, n$ and $c^i$ constant and $|c^i|$ sufficiently small are also submanifolds, and they provide a foliation of $M$: a set of submanifolds such that every point in $M$ is on precisely one of them.

Given $n$ vector fields on $M$, what are the conditions that there exists a foliation of $M$ by $n$-dimensional hypersurfaces such that its tangent space at each point in $M$ is spanned by those vector fields? The answer is

Frobenius’ Theorem (vector field version): Let $\{V_\alpha\}, \alpha = 1, \ldots, n$ be vector fields on $M$. Their integral curves mesh to form a foliation of $M$ by submanifolds if and only if the commutator of any two of these vector fields is a linear combination of all these vector fields,

$$[V_\alpha, V_\beta] = \sum_\gamma c_{\alpha\beta}^\gamma V_\gamma. \quad (118)$$

This is a generalisation of the theorem that we can find coordinates such $V = \partial/\partial x^i$ to $n$ from one vector field to $n$ commuting vector fields. A proof, by induction, is given in Schutz or in Walker. Intuitively, the condition means that as we Lie-drag any one of these vector fields along any other, we stay in the surface. Another version of Frobenius’ Theorem and its proof use the machinery of differential forms.

### 2.10 Application: axisymmetry

Lie derivatives are closely related to symmetries. As an example, consider a fluid flow that is restricted to be axisymmetric. In spherical polar coordinates adapted to the symmetry, this just means that the density $\rho$ and velocity $v^i$ are independent of the angle $\varphi$, or $\rho = \rho(r, \theta)$ and $v^i = v^i(r, \theta)$ for $i = r, \theta, \varphi$. The vector

$$X = \frac{\partial}{\partial \varphi} \quad (119)$$

is said to be the **generator of a symmetry**, in the sense that

$$\mathcal{L}_X \rho = \frac{\partial \rho}{\partial \varphi} = 0, \quad (120)$$

$$\mathcal{L}_X v^i = \frac{\partial v^i}{\partial \varphi} = 0, \quad i = r, \theta, \varphi \quad (121)$$

and $\rho$ and $v$ are said to be **invariant** under (the symmetry generated by) $X$. 

---

22
As these equations are tensor equations, we can express them in arbitrary coordinates. From the chain rule
\[ X^i = \frac{\partial x^i}{\partial \varphi}, \]  
(122)

and from our formulae above
\[ \mathcal{L}_X \rho = X^i \rho_{,i} = 0, \]  
(123)
\[ \mathcal{L}_X v^i = X^j v^i_{,j} - v^j X^i_{,j} = 0. \]  
(124)

In Cartesian coordinates \((x, y, z) = (r \cos \theta \cos \varphi, r \cos \theta \sin \varphi, r \sin \theta)\), we have
\[ X = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}. \]  
(125)

The partial derivative terms are expected, but the lower order terms are not. They can be thought of as the rotation of the basis vectors in the \(x\) and \(y\) direction as we rotate around the \(z\) axis.

2.11 Exercises

1. Show that (77) and (78) do not transform as tensors.

2. Derive (92).

3. Consider the composition of two 1-parameter family of maps of \(M\) to itself, and calculate in coordinates
\[ V = \frac{\partial}{\partial t} \psi_t (f_t(p)) \bigg|_{t=0}. \]  
(129)

4. Show that, acting on scalars and vector fields, the commutator of two Lie derivatives is the Lie derivative along the commutator of vector fields, or
\[ [\mathcal{L}_V, \mathcal{L}_W] = \mathcal{L}_{[V, W]}. \]  
(130)

5. Assume that in a coordinate basis the Lie derivative is given by (107). Show that in an arbitrary basis \(e_a\) of vector fields the more general expression
\[ (\mathcal{L}_V U)^a = V^b e_b U^a - U^b e_b V^a + V^b U^c (\mathcal{L}_{e_b} e_c)^a \]  
(131)
applies. Hint: use the product rule on \(U = U^a e_a\), where \(U^a\) is a function and \(e_a\) is a vector field.
6. Show that if $V$ and $W$ are linear combinations (not necessarily with constant coefficients) of $m$ vector fields that commute with each other, then $[V, W]$ is a linear combination of the same $m$ vector fields (use the previous question).

b) Prove the same result when the $m$ vector fields have Lie brackets which are nonvanishing linear combinations of the $m$ vector fields.

7. The vector fields $l_x, l_y$ and $l_z$ generate rotations about the $x, y$ and $z$ axes and obey $[l_x, l_y] = l_z$ and cyclic. Define $L^2 := L^2_{x} + L^2_{y} + L^2_{z}$.

a) Show that $L_{l_z}$ and $L^2$ commute.

b) Show for a scalar $\phi$ that

$$L^2\phi = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \phi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \phi}{\partial \varphi^2}$$

(132)

8. Show that if $T$ is invariant under $V$ and $W$, it is invariant under $aV + bW$ where $a$ and $b$ are constants.
3 Linear connections and curvature

3.1 Linear connections

The Lie derivative gives us a notion of the derivative of a tensor along a vector field. However, \( \mathcal{L}_X T \) contains derivatives of \( X \) as well as of \( T \). In other words, it depends on the vector field \( X \) and not just on a vector \( X \) at one point.

The covariant derivative does not have this disadvantage, but it requires an additional (non-unique) structure on the manifold called a linear connection.

**Definition:** A linear connection on \( M \) is a map from any two smooth vector fields \( X \) and \( Y \) to a vector field \( \nabla_X Y \) that is linear in \( X \) and acts as a derivative operator on \( Y \), that is

\[
\nabla_{\phi X + Y} Z = \phi \nabla_X Z + \nabla_Y Z, \tag{133}
\]

\[
\nabla_X (\phi Y + Z) = \phi \nabla_X Y + (X\phi) Y + \nabla_X Z. \tag{134}
\]

for vector fields \( X, Y, Z \) and a scalar field \( \phi \).

Note that because \( \nabla_X Y \) is not linear in \( Y \) (instead we have the product rule (134)) \( \nabla \) does not define a \((1,2)\)-tensor, but

\[
\nabla Y \tag{135}
\]

does define a \((1,1)\) tensor, the covariant derivative of \( Y \), which obeys

\[
(\nabla Y)X = \nabla_X Y, \tag{136}
\]
or in abstract index notation

\[
(\nabla_X Y)^i = X^j (\nabla Y)^i_j. \tag{137}
\]

More commonly we write this \((1,1)\)-tensor in one of two ways as

\[
(\nabla Y)^{i}_j =: \nabla_j Y^i =: Y^{i}_{;j}. \tag{138}
\]

Note that both these notations are potentially confusing: \( \nabla_j Y^i \) or \( Y^{i}_{;j} \) are the components of a single \((1,1)\)-tensor (or in abstract index notation represent a \((1,1)\)-tensor).

For any vector \( X \) and scalar field \( \phi \) we define

\[
\nabla_X \phi := X\phi \tag{139}
\]

In terms of an arbitrary basis \( \{e_a\} \) of vector fields, keeping in mind that the components \( X^a \) in a *fixed* basis are functions, we have

\[
\nabla_{(X^a e_a)} (Y^b e_b) = X^a \nabla_{e_a} (Y^b e_b) = X^a [(e_a Y^b) e_b + Y^b (\nabla_{e_a} e_b)] = X^a [e_a Y^c + Y^b \Gamma^c_{ab}] e_c, \tag{140}
\]

where the connection coefficients

\[
\nabla_a e_b := \nabla_{e_a} e_b := \Gamma^c_{ab} e_c \tag{141}
\]
give the covariant derivatives of one basis vector along another.
We can now calculate the transformation law for the $\Gamma^c_{ab}$ under a change of basis. We shall do this in a coordinate basis. The connection coefficients in a coordinate basis are called Christoffel symbols. From (140)\[
abla_i Y^j := Y^j,_{i} + \Gamma^j_{ik} Y^k.\]This must transform as a tensor under coordinate changes, which allows us to read off the transformation law of the Christoffel symbols. Under a change of coordinates $x^i \mapsto \tilde{x}^p$,\[
abla_i Y^j = Y^j,_{i} + \Gamma^j_{ik} Y^k = \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial x^j}{\partial \tilde{x}^q} \left( Y^q,_{p} + \Gamma^q_{pr} Y^r \right) = \frac{\partial x^p}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^j}{\partial x^i} \left( \frac{\partial x^q}{\partial \tilde{x}^l} Y^l,_{m} + \frac{\partial^2 x^q}{\partial \tilde{x}^l \partial \tilde{x}^m} Y^l \right) \frac{\partial x^m}{\partial \tilde{x}^p} + \Gamma^q_{pr} \frac{\partial x^r}{\partial \tilde{x}^k} Y^k = Y^j,_{i} + \left( \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^r}{\partial x^i} \frac{\partial x^p}{\partial \tilde{x}^l} \Gamma^q_{pr} \frac{\partial x^m}{\partial \tilde{x}^p} + \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial^2 x^q}{\partial \tilde{x}^l \partial \tilde{x}^m} \right) Y^k\](143)and so we must have\[
\Gamma^j_{ik} = \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^r}{\partial x^i} \frac{\partial x^p}{\partial \tilde{x}^l} \Gamma^q_{pr} \frac{\partial x^m}{\partial \tilde{x}^p} + \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial^2 x^q}{\partial \tilde{x}^l \partial \tilde{x}^m} \]\(144)Note the first term looks as if the connection coefficient transforms as a $(1,2)$-tensor, but the second term is inhomogeneous.

In order to define the covariant derivative of a tensor of arbitrary type, we define $\nabla$ to obey the product rule (as we did for the Lie derivative):\[
\nabla_X (T \cdot S) = (\nabla_X T) \cdot S + T \cdot (\nabla_X S),\]
(145)where $T$ and $S$ are for any two tensors, perhaps contracted. From these two rules, applied to $f = T^i S_i$ one finds\[
\nabla_i \omega_j = \omega_{j,i} - \Gamma^k_{ij} \omega_k,\]
(146)and hence the general rule\[
\nabla_i T^{j\ldots k} = T^{j\ldots k}_{\quad i} + \Gamma^j_{il} T^{l\ldots k} + \cdots - \Gamma^j_{ik} T^{j\ldots l} - \cdots\]
(147)The dots represent one Christoffel term with a plus sign for each extra up index on $T$, and one Christoffel term with a minus sign for each extra down index on $T$. (Note: In these notes, the “derivative index” is always the middle index on the connection coefficient, as in the books Aubin and CD/DM/DB and most other pure mathematics texts. In Schutz, and most texts on relativity, the derivative index is the last index.)

### 3.2 Parallel transport and geodesics

We have already mentioned that, in contrast to Euclidean geometry, on a manifold we have no immediate way of comparing tangent vectors at different points.
Let us try to move a vector $Y$ defined at a point to another point along an integral curve of the vector field $X$. With $t$ the parameter along any such integral curve, we have

$$Xf = X^i f_i = \frac{df}{dt}. \quad (148)$$

We can **Lie drag** the vector field $Y$ along the vector field $X$ by defining

$$\mathcal{L}_X Y = 0. \quad (149)$$

In coordinates this gives

$$\frac{dY^i}{dt} - Y^j X^i_{,j} = 0 \quad (150)$$

for $Y^i(t)$ along the curve. To solve this as an ordinary differential equation, we need to know $X$ in a neighbourhood of the integral curve, not just on the integral curve itself. In other words, we must define $X$ as a vector field, not just as the tangent vector to a single curve. Intuitively, this must be so because Lie dragging of a vector transports an infinitesimal piece of a curve to which $Y$ is tangent (the “tip” and “tail” of $Y$) along infinitesimally neighbouring integral curves of $X$.

Alternatively, we can **parallelly transport** $X$ along $Y$ with the help of the linear connection $\nabla$ by defining

$$\nabla_X Y = 0. \quad (151)$$

In coordinates, this is

$$\frac{D}{Dt} Y^i := \frac{dY^i}{dt} + \Gamma^i_{jk} X^j Y^k = 0, \quad (152)$$

where $t$ is the parameter of the integral curve of $X$. Note this only requires $X$ and $Y$ along the integral curve itself. This means that we can define $Y^i$ at a point $p$, define $X$ as the tangent vector of a single curve $c(t)$, and parallelly transport $X$ to any point on $c$.

Finally, we want the idea of “going in a straight line”. We cannot Lie-drag a vector field along itself, but we can parallelly transport a vector field along itself. We define a **geodesic** of the connection $\nabla$ as the integral curve of a vector field $X$ that obeys

$$\nabla_X X = 0. \quad (153)$$

In coordinates this gives

$$\frac{dX^i}{dt} + \Gamma^i_{jk} X^j X^k = 0. \quad (154)$$

We see that in fact we do not need to define $X$ as a vector field. Rather, any point $p$ and vector $X_p$ at $p$ locally define a geodesic curve through $p$ with tangent vector $X_p$ at $p$. To see this, note that in coordinates, along the curve

$$\frac{dx^i}{dt} = X^i, \quad (155)$$

and so we have the system of ordinary differential equations

$$\frac{d^2 x^i}{dt^2} + \Gamma^i_{jk}(x^j(t)) \frac{dx^j}{dt} \frac{dx^k}{dt} = 0, \quad x^i(0) = x^i_p, \quad \frac{dx^i}{dt}(0) = X^i_p, \quad (156)$$

which has a unique solution for small $t$. 

27
3.3 The exponential map

**Definition:** Given a manifold $M$ with a connection $\nabla$, and a point $p$, the **exponential map** is defined to be the map from a tangent vector $X$ at $p$ to the point $q$ a parameter distance of 1 along the geodesic through $p$ with tangent vector $x$.

**Theorem:** This map is locally one-to-one.

**Idea of proof:** If $x_p^i$ are the coordinates of $p$ in a chart and $X_p^i$ the components of $X_p$ in the same chart, an expansion of the solution of (156) in powers of $t$ gives

$$x^i(t) = x_p^i + tX_p^i + O(t^2).$$

So the map from $X_p^i$ to $x^i(t)$ is invertible for small enough $t$. But the solution is invariant under the rescaling

$$t \to \alpha^{-1}t, \quad X_p^i \to \alpha X_p^i,$$

meaning that the geodesic with initial velocity $\alpha X$ reaches the same point at time $\alpha^{-1}$. Hence the map from $X_p^i$ to $x^i(1)$ is also invertible for small enough $X_p^i$.

**Corollary:** There is a unique geodesic through any two points on $M$ that are sufficiently close.

**Idea of proof:** Given $p$ and $q$, we have $X_p$. Then $p$ and $X_p$ define a geodesic.

**Corollary and definition:** In a sufficiently small neighbourhood of $p$, the components of $X_p$ are coordinates on $M$. They are called **normal coordinates** at $p$.

Consider the geodesic from $p$ to $q$, and a basis $e_a$ of $T_pM$ so that $X_p = X_p^a e_a$.

The normal coordinates of $q$ are then $x^a = X_p^a$.

**Theorem:** In normal coordinates, $\Gamma^a_{(bc)} = 0$ at $p$.

This is an important tool for proofs by calculation. It means that we can always find coordinates such that $\Gamma_{(jk)} = 0$ vanishes at a given point (that is, $\Gamma_{jk} = 0$ for a torsion-free connection). Note that it will not vanish at any other point, nor do its derivatives vanish at $p$.

**Idea of proof:** Because of the rescaling invariance, the normal coordinates of any point on the geodesic must be proportional to $X_p^a$. Hence the geodesic through $p$ with tangent vector $X$ is given in the normal coordinates by $x^a(t) = s(t)X_p^a$, with $s(1) = 1$ and $ds/dt \neq 0$. Hence

$$\frac{d^2s}{dt^2}X_p^a + \left(\frac{ds}{dt}\right)^2 \Gamma^a_{bc}X_p^bX_p^c = 0.$$ (159)

But at $p$, we could have chosen any $X_p^a$, and so at $p$, (159) must hold for any $X_p^a$. This can be true only if both $\Gamma^a_{(bc)} = 0$ and $d^2s/dt^2 = 0$ at $p$. (However, once we have moved away from $p$, we have chosen our geodesic, and $X_p^a$ can no longer be considered arbitrary, so this result holds only at $p$.)

3.4 Torsion

**Definition and theorem:** The map defined by

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]$$ (160)
defines a (1,2)-tensor, the torsion tensor.

**Proof:** Clearly $T(X,Y)$ is a vector field, and so we only need to check linearity. As $T(X,Y) = -T(Y,X)$, we only need to check linearity in one argument. Linearity under addition is obvious. To check linearity under multiplication by a scalar, we calculate

$$
T(\phi X, Y) = \nabla_{\phi X} Y - \nabla_Y (\phi X) - [\phi X, Y] \\
= \phi \nabla_X Y - (\phi \nabla_Y X + (Y \phi) X) - (\phi [X, Y] - (Y \phi) X) \\
= \phi (\nabla_X Y - \nabla_Y X - [X, Y]) \\
= \phi T(X, Y).
$$

(A161)

A connection with vanishing torsion tensor is called torsion-free or symmetric.

**Corollary:** From (140) and (52), the components of the torsion tensor are given by

$$
T^c_{ab} = 2 \Gamma^c_{[ab]} - C^c_{ab}.
$$

(A162)

In a coordinate basis, where the structure constants $C^i_{jk}$ vanish, the torsion tensor is just given by the antisymmetric part of the linear connection,

$$
T^i_{jk} = 2 \Gamma^i_{[jk]}
$$

(A163)

in a coordinate basis, which explains the term “symmetric connection”.

In a coordinate basis, for a scalar field $\phi$,

$$
\nabla_i \nabla_j \phi = \nabla_i \phi, j - \Gamma^k_{ij} \phi, k
$$

(A164)

and so, again in a coordinate basis,

$$
(\nabla_i \nabla_j - \nabla_j \nabla_i)\phi = -2 \Gamma^k_{[ij]} \phi, k = -T^k_{ij} \nabla_k \phi.
$$

(A165)

But the left-hand side and right-hand side of this are tensors, and so this holds in any basis,

$$
(\nabla_a \nabla_b - \nabla_b \nabla_a + T^c_{ab} \nabla_c)\phi = 0.
$$

(A166)

We can also write the Lie bracket using covariant derivatives and the torsion tensor. We have

$$
[X, Y]^i = X^j Y^i_{,j} - Y^j X^i_{,j} = X^j Y^i_{,j} - Y^j X^i_{,j} - 2 \Gamma^i_{jk} X^j Y^k
$$

(A167)

in a coordinate basis, and therefore

$$
[X, Y]^a = X^b \nabla_b Y^a - Y^b \nabla_b X^a - T^a_{bc} X^b Y^c
$$

(A168)

in a general basis and in abstract index notation.

### 3.5 Curvature

**Definition and theorem:** The map defined by

$$
R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.
$$

(A169)

defines a (1,3)-tensor, the curvature tensor of the linear connection $\nabla$. 

29
Proof: Clearly \( R(X,Y)Z \) is a vector field, so we only need to check linearity in \( X, Y \) and \( Z \). Moreover, \( R(X,Y)Z = -R(Y,X)Z \), and linearity under addition is obvious, and so we only need to check linearity in \( X \) and \( Z \) under multiplication by scalars. We calculate

\[
R(\phi X, Y)Z = \nabla_{\phi X} \nabla_Y Z - \nabla_Y \nabla_{\phi X} Z - \nabla_{[\phi X,Y]} Z \tag{170}
\]

\[
= \phi \nabla_X \nabla_Y Z - \nabla_Y (\phi \nabla_X Z) - \nabla_{\phi [X,Y]} (-\nabla_Y \phi X)Z \tag{171}
\]

\[
= \phi \nabla_X \nabla_Y Z - \phi \nabla_Y \nabla_X Z - (Y\phi) \nabla_X Z - (Y\phi) \nabla_X Z \tag{172}
\]

\[
= \phi R(X,Y)Z \tag{173}
\]

and

\[
R(X,Y)(\phi Z) = \nabla_X \nabla_Y (\phi Z) - \nabla_Y \nabla_X (\phi Z) - \nabla_{[X,Y]} (\phi Z) \tag{174}
\]

\[
= \nabla_X (\phi \nabla_Y Z + (Y\phi)Z) - (X \leftrightarrow Y) - \phi \nabla_{[X,Y]} Z - ([X,Y] \phi)Z \tag{175}
\]

\[
= \phi \nabla_X \nabla_Y Z + (X \phi) \nabla_Y Z + (X(Y\phi))Z + (Y\phi) \nabla_X Z - (X \leftrightarrow Y) - \phi \nabla_{[X,Y]} Z - (X(Y\phi) - Y(X\phi))Z \tag{176}
\]

\[
= \phi R(X,Y)Z. \tag{177}
\]

We now calculate the components of the curvature tensor in an arbitrary basis:

\[
R(e_a, e_b)e_c = \nabla_a \nabla_b e_c - \nabla_b \nabla_a e_c - \nabla_{[e_a , e_b]} e_c \tag{178}
\]

\[
= \nabla_a (\Gamma^f_{b c} e_f) - (a \leftrightarrow b) - [e_a , e_b]^d \nabla_d e_c \tag{179}
\]

\[
= (e_a \Gamma^f_{b c})e_f + \Gamma^d_{b c} \Gamma^f_{a d} e_f - (a \leftrightarrow b) - C^d_{a b} \Gamma^f_{d c} e_f \tag{180}
\]

\[
=: R^f_{c a b} e_f \tag{181}
\]

where

\[
R^f_{c a b} = e_a \Gamma^f_{b c} + \Gamma^d_{b c} \Gamma^f_{a d} - (a \leftrightarrow b) - C^d_{a b} \Gamma^f_{d c} \tag{182}
\]

In a coordinate basis only,

\[
R^f_{i j} = \Gamma^k_{i j} e_k + \Gamma^m_{j i} e_m - (i \leftrightarrow j), \tag{183}
\]

as the structure constants of a coordinate basis vanish.

In terms of an arbitrary basis, or in abstract index notation, we have, using (168),

\[
X^a Y^b Z^c R^d_{c a b} = X^a \nabla_a (Y^b \nabla_b Z^d) - Y^a \nabla_a (X^b \nabla_b Z^d) - (X^b \nabla_b Y^c - Y^b \nabla_b Y^c - X^a Y^b T^c_{a b}) \nabla_c Z^d \tag{184}
\]

\[
= X^a Y^b \left[ (\nabla_a \nabla_b - \nabla_b \nabla_a) Z^d + T^c_{a b} \nabla_c Z^d \right] \tag{185}
\]

and so

\[
(\nabla_a \nabla_b - \nabla_b \nabla_a + T^c_{a b} \nabla_c) Z^d = R^d_{c a b} Z^c, \tag{186}
\]

the Ricci identity.

Applying (166) to the scalar \( \omega_a Z^a \) and using (186), we obtain the Ricci identity for covectors,

\[
(\nabla_a \nabla_b - \nabla_b \nabla_a + T^c_{a b} \nabla_c) \omega_d = -R^c_{d a b} \omega_c. \tag{187}
\]
For any torsion-free connection, the first and second Bianchi identities
\[ R^a_{\ [bcd]} = 0, \]  
\[ \nabla_{[e} R^a_{\ b]cd} = 0 \]  
hold. (See Aubin for the version with torsion). The notation in (189) means “antisymmetrise over ecd”. Both identities can be proved by calculation in normal coordinates, using that \( \Gamma_{ijk} = 0 \) at one point. As they are tensor equations, they must hold in all coordinates. They also follow from the covector form of the Ricci identity, using the exterior derivative (see later).

3.6 Flat connections

A connection (or covariant derivative) with \( R^a_{\ bcd} = 0 \) is called flat. For example, if \( \nabla_i = \partial/\partial x^i \) in some coordinate system, then the curvature tensor vanishes in that coordinate system because the Christoffels vanish in that coordinate system, and so this connection is flat. We state without proof that the reverse also holds:

**Theorem:** The curvature and torsion vanish if and only if there is a coordinate chart in which the connection coefficients vanish.

The Christoffel symbol \( \Gamma_{ijk} \) is not a tensor because under a change of coordinates the second term in (144) is not linear in \( \Gamma_{ijk} \). But this second, inhomogeneous term is actually independent of \( \Gamma_{ijk} \) and depends only on the old and new coordinates. Hence it is easy to see (exercise) that the difference between the Christoffel symbols of two connections transforms as a tensor. Hence, if \( \nabla \) and \( \tilde{\nabla} \) are two different connections, we can write in abstract index notation
\[ \nabla_a X^b - \tilde{\nabla}_a X^b = D^b_{\ ac} X^c \]  
where
\[ D^b_{\ ac} := \Gamma^b_{\ ac} - \tilde{\Gamma}^b_{\ ac} \]
is a (1,2)-tensor.

Now, given any chart \( \tilde{\phi} \), we define the flat connection \( \nabla \) associated with \( \tilde{\phi} \) by defining
\[ \nabla_i X^j = X^j,_{i} \iff \tilde{\Gamma}^i_{\ jk} = 0 \]  
in the chart \( \tilde{\phi} \) only. (192)

Because the transformation of the Christoffel symbols is inhomogeneous, \( \Gamma^i_{\ jk} \) will in general not vanish in any other chart.

In a general chart \( \phi \) we have
\[ \nabla_a X^b = X^j,_{ai} + \Gamma^j_{\ i} X^k. \]  
Here the left-hand side represents the components of a (1,1)-tensor, and so does the right-hand side, but neither term on the right-hand side is a tensor. However, we can write
\[ \nabla_a X^b = \tilde{\nabla}_a X^b + D^b_{\ ac} X^c, \]
where each of the three terms is a tensor, while in the preferred chart \( \tilde{\phi} \) this equation reduces to (193). Therefore, we can always think of the partial derivative in a particular coordinate system as a covariant derivative with a flat connection. Conversely, for any flat connection we can locally find coordinates such that the covariant derivative reduces to the partial derivative in those coordinates.
3.7 Geodesic deviation

Let

\[ c : \mathbb{R} \times \mathbb{R} \ni (t, s) \mapsto c(t, s) \quad (195) \]

be a \( C^\infty \) map, and define the vector fields

\[ X := \frac{\partial}{\partial t}, \quad Y := \frac{\partial}{\partial s}. \quad (196) \]

\( t \) and \( s \) are coordinates on a 2-dimensional submanifold, so that

\[ [X, Y] = 0. \quad (197) \]

Let the curve \( c_s(t) = c(t, s) \) for fixed \( s \) be a geodesic with respect to a connection \( \nabla \) on \( M \),

\[ \nabla_X X = 0. \quad (198) \]

The vector \( Y \) intuitively represents the separation of neighbouring curves \( c_s(t) \) at the same \( t \), and is called the deviation vector or rigging vector.

Let \( \nabla \) be torsion-free, so that

\[ \nabla_X Y - \nabla_Y X = [X, Y] = 0. \quad (199) \]

We now have

\[ \nabla_X \nabla_Y X = \nabla_X (\nabla_Y X) = (\nabla_X \nabla_Y X - \nabla_Y \nabla_X X) = R(X, Y)X \quad (200) \]

or, using the notation \( D/Dt \) for parallel transport along \( X \),

\[ \frac{D^2}{Dt^2} Y^i = R^i_{jkl} X^j X^k Y^{l}. \quad (201) \]

This is the geodesic deviation equation. Intuitively, we can think of \( \nabla_X \nabla_Y X \) as the “relative acceleration” (with respect to the parameter \( t \)) between the neighbouring geodesics. In particular, if \( \nabla_X Y = 0 \) at \( t = 0 \), that is, neighbouring geodesics are initially parallel, but they will not stay parallel in general. Considering now any \( X \) and \( Y \), we see that any two initially parallel neighbouring geodesics will remain parallel if and only if the Riemann tensor vanishes identically.

3.8 The metric-covariant derivative

If the manifold \( M \) is equipped with a metric, there is a preferred covariant derivative, the metric-covariant derivative or Riemannian connection, which is defined by the properties

\[ T^a_{bc} = 0 \quad (202) \]

(the connection is symmetric), and

\[ \nabla_a g_{bc} = 0 \quad (203) \]
It can be shown (see exercise) that its Christoffel symbols are
\[ \Gamma^i_{jk} = \frac{1}{2} g^{il} \left( -g_{jk,l} + g_{kl,j} + g_{jl,k} \right). \]
(Note \( \Gamma^i_{jk} = \Gamma^i_{kj} \) as required).

A prime example is Euclidean space \( \mathbb{R}^n \). We have seen that in Cartesian coordinates the Euclidean metric \( \gamma_{ab} \) on \( \mathbb{R}^n \) has components \( \gamma_{ij} = \delta_{ij} \), and (204) then gives us \( \Gamma^i_{jk} = 0 \). Hence in Cartesian coordinates the covariant derivative is simply the partial derivative, or \( \nabla_i = \partial / \partial x^i \). As \( \delta_{ij} \) is constant, \( \nabla_i \gamma_{jk} = 0 \) is clearly obeyed in Cartesian coordinates. It is however true in all coordinates, even if the \( \Gamma^i_{jk} \) do not vanish (see exercise).

The curvature tensor of a metric connection is also called the **Riemann tensor**, and it contains information about the geometry, or shape, given to the manifold by that metric. A metric with vanishing Riemann tensor is again called **flat**. The Riemann tensor associated with a metric-covariant derivative, with its first index lowered with \( g_{ab} \), obeys the identity
\[ R_{abcd} = R_{cdab}, \]
in addition to the \( R^a_{bcd} = R^a_{bdc} \) and the Bianchi identities. We define the **Ricci tensor**
\[ R_{ab} := R_{acb}^\ c \]
and **Ricci scalar**
\[ R := R^a_a \]
where we have used the convention of moving indices implicitly with the metric, and hence the **Einstein tensor**
\[ G_{ab} := R_{ab} - \frac{1}{2} \gamma_{ab} R. \]

We can also derive the **contracted Bianchi identities**
\[ \nabla_a G^{ab} = 0. \]

To motivate the name Einstein tensor, we mention in passing that in **general relativity**, spacetime is a 4-dimensional manifold with a pseudo-Riemannian metric \( g_{ab} \) (three positive eigenvalues for space and one negative eigenvalue for time), and the Einstein equations are \( G_{ab} = 8 \pi T_{ab} \) where the stress-energy tensor \( T_{ab} \) represent the matter. The contracted Bianchi identities then give \( \nabla_a T^{ab} = 0 \), which governs the behaviour of the matter. Also, point particles move on geodesics of \( g_{ab} \), so that the spacetime metric in some sense replaces the Newtonian gravitational field. Spacetimes without matter obey \( G_{ab} = 0 \) and are called Einstein spaces in the mathematics literature.

### 3.9 Outlook: connections on fibre bundles

This subsection is very, very handwaving, and is intended mainly to establish connections with topics you may have learned about elsewhere, such as quantum field theory or Lie groups. See Sections II.B.2 and Vbis of CB/DM/DB for proper definitions, or Isham for a physics point of view.
A fibre bundle is a manifold $E$ that is locally the product of a base manifold $M$ and a fibre manifold $F$, together with a projection map $\pi$ from a point in the bundle to a point in the base (which fibre are we on?) and a map from a sample fibre $F$ to $E$ (where on the fibre are we?). We can write this as

$$F \to E \to M$$

(210)

An example for a fibre bundle that is not globally a product is the Möbius strip, with bundle $S^1$ and fibre $\mathbb{R}^1$.

A vector bundle is a fibre bundle where the fibre manifold is a vector space. An example is the tangent bundle of a manifold, where the base is the manifold and the fibre its tangent space at a point. The tangent bundle of $S^2$, for example, is locally but not globally the product manifold $S^2 \times \mathbb{R}^2$.

A Lie group $G$ is a group that is also a manifold. An element $g$ of the group then gives a map $f_g$ from $G$ to itself, which acts by $f_g h = gh$ (left translation). A left-invariant vector field $V$ on $G$ is a vector field that is invariant under this group action in the sense that

$$(f_g)_* V(h) = V(gh).$$

(211)

It turns out that if $G$ is $n$-dimensional (as a manifold), then the left-invariant vector fields form a vector space also of dimension $n$, and hence there is an obvious identification of this space with $T_e G$, the tangent space of the Lie group at its unit element. We call this vector space the Lie algebra of $G$.

A principal fibre bundle is a fibre bundle acted upon freely by a Lie group $G$, and such that the base manifold $M$ is isomorphic to $E/G$, and the fibre is isomorphic to $G$:

$$F \simeq G \to E \to M \simeq E/G$$

(212)

An example is the frame bundle of a manifold, where the fibre is the set of frames. A frame at a point is a specific ordered basis of the tangent space at that point. With respect to some fixed frame $\{e_a\}$, any other frame $\{\tilde{e}_a\}$ can be specified as

$$e_a = M_{a}^{\ b} \tilde{e}_b,$$

(213)

where $M_{a}^{\ b} \in GL(n, \mathbb{R})$, and changes of basis are expressed by multiplying $M_{a}^{\ b}$ by some $N_{a}^{\ b} \in GL(n, \mathbb{R})$. So $GL(n, \mathbb{R})$ is both the fibre (a frame) and the group acting on it (a change of frame).

A cross-section through a fibre bundle $E$ is a generalisation of a “field” or “function” on its base manifold. For example, a cross section through the tangent bundle defines vector field on the base manifold. But a cross-section through the Möbius strip is only locally a function on $S^1$, while globally it is double-valued.

We have made these definitions to look at the linear connection in a more general light. A connection on a fibre bundle identifies neighbouring fibres. An abstract point of view is that a connection provides a split of the $TE$, the tangent space to the fibre bundle, into a product of a horizontal vector space (with dimension of $M$) and a vertical vector space (with dimension of $F$). This can be done by providing a connection 1-form $\omega$ on $E$ such that a tangent vector $X$ is horizontal if and only if $\omega(X) = 0$.

However, a more practical point of view is to introduce a local representation of the connection 1-form. On a vector bundle, we can write this as
\[ \Gamma^{\alpha \beta}_{\gamma}, \] where \( \alpha \) is an index on the tangent space \( TM \) of the base manifold \( M \), and \( \alpha \) and \( \beta \) are indices on the fibre (vector) space (which need not even have the same dimension as \( TM \)). We can then define a **covariant derivative** on a vector bundle acting on a vector field \( \xi \) as

\[
\nabla_i \xi^\alpha = \xi^\alpha_i + \Gamma^{\alpha \beta}_{\gamma} \xi^\beta.
\]

From there, we can define a curvature tensor of this connection written as \( R^{\alpha \beta}_{\gamma \delta} \).

For a principal fibre bundle, this point of view gives us a local representation of the connection 1-form as a 1-form in the base manifold that takes values in the Lie algebra of the fibre group. This is also called a **Yang-Mills field** in a physics context. We can write it as \( A_i \) or, if we wanted to show that it is Lie algebra-valued, as

\[ A_i = A^A_i e_A, \]

where \( \{e_A\} \) is a basis of the Lie algebra. If \( \psi \) is an element of the Lie group, we can write the corresponding covariant derivative as

\[
\nabla_i \psi = \psi_i + A_i. \]

The corresponding curvature 2-form is written as \( F_{ij} \), or \( F_{ab} \) in abstract index notation, antisymmetric in \( a \) and \( b \), and taking again values in the Lie algebra.

In the particular case where the Lie group is \( U(1) \), its Lie algebra is \( \mathbb{R} \), and in this case \( F_{ab} \) takes values in \( \mathbb{R} \), in other words, it is just an antisymmetric \((0,2)\)-tensor. In the context of electromagnetism, this tensor is called the **field strength tensor**, or **Faraday tensor**. In the context of gauge theory in particle physics, where the Lie group is for example \( SU(3) \) to describe the strong interaction, this object is called the field strength tensor or the curvature tensor.

We loop back to the familiar linear connection on a manifold. It is clear that this is a connection on a vector bundle, namely the tangent bundle. But if we use a non-coordinate basis, it becomes a connection on the frame bundle, which, as we have seen, is an example of a principal fibre bundle, with group \( GL(n, \mathbb{R}) \). Now the Lie algebra of \( GL(n, \mathbb{R}) \) is also \( GL(n, \mathbb{R}) \), and is therefore \( n^2 \) dimensional. We can then interpret the indices \( a \) and \( b \) on \( \Gamma^{\alpha \beta}_{\gamma \delta} \) as indices on the fibre space of the tangent bundle, as in (214), or, alternatively, the pair of indices \( a \) and \( b \) as a composite index \( A \) on the Lie algebra \( GL(n, \mathbb{R}) \) of the frame bundle, as in (215).

To recap all this potentially confusing (but fairly standard) notation, the most abstract expression for the curvature tensor of a connection on a principal fibre bundle is \( F_{ab} \), taking values in the Lie algebra of the fibre group, with \( ab \) an antisymmetric pair of (abstract) indices in the tangent space of the base manifold. Introducing an explicit basis \( \{e_A\} \) of the Lie algebra, we can write this as \( F_{ab} = F^A_{ab} e_A \). If the Lie group is a matrix group, we can write this basis as \( \{e_A^\alpha \beta\} \), where \( \alpha \) and \( \beta \) are matrix indices. Hence we can also write \( F^{\alpha \beta}_{ab} = F^A_{ab} e^A^{\alpha \beta} \). Finally, if the fibre is \( TM \) itself, the indices \( \alpha \) and \( \beta \) become indices in \( TM \), and we arrive at \( \mathbf{F}_{dab} \), which is now called the curvature or Riemann tensor rather than the field strength tensor, and written as \( R^c_{dab} \).

### 3.10 Exercises

1. Derive (147), for the special cases of a \((1,1)\)-tensor (or for the general case if you are good with notation).
2. Show that in (108) the partial derivative can be replaced by any covariant derivative, because the connection coefficients cancel out.

3. Let \( x^i \) represent the Cartesian coordinates \((x, y, z)\), and let \( x^i \) represent the spherical polar coordinates \((r, \theta, \phi)\). Assume that in Cartesian coordinates \( \tilde{\gamma}_{ij} = \delta_{ij} \). Use (35) to calculate \( \gamma_{ij} \) and hence \( \gamma^{ij} \) in spherical polar coordinates. Calculate the \( \Gamma^i_{jk} \) in spherical polar coordinates using (204).

4. Assume that in Cartesian coordinates \( \tilde{\Gamma}^i_{jk} = 0 \), and use (144) to calculate the \( \Gamma^i_{jk} \) in spherical polar coordinates. You should get the same answer as before.

5. Use \( \gamma^{ij} \) and \( \Gamma^i_{jk} \) from the previous questions to write down the three components of Euler’s equation

\[
\dot{v}^i + v^j \nabla_j v^i = -\frac{1}{\rho} \nabla^i p
\]

in spherical polar coordinates. (The time derivative \( \partial / \partial t \) denoted by an overdot does not change under a change of spatial coordinates.)

6. Assume that \( \Gamma^i_{jk} = 0 \) is symmetric in \( j \) and \( k \) (no torsion). Define \( \Gamma^{ijk} := g_{il} \Gamma^l_{jk} \) as a shorthand notation. Write out \( \nabla_i g_{jk} = 0 \) in full using the connection coefficients \( \Gamma \). By permuting indices, also write down \( \nabla_j g_{ik} = 0 \) and \( \nabla_k g_{ij} = 0 \). Add these three equations (perhaps with a minus sign) to solve for \( \Gamma^i_{jk} \). Raise an index to obtain \( \Gamma^i_{jk} \).

7. Calculate, in coordinates, the components of the torsion

\[
\nabla_U V - \nabla_V U - [U, V].
\]

Your answer should be proportional to \( U \) and \( V \) (undifferentiated) and an object for you to define, the torsion tensor, which is made from the connection coefficients.

8. Show that the difference between two Christoffel symbols transforms as a tensor. Hence the difference of two connections is a tensor.

9. From the definition (187), by expanding \( \nabla_i \nabla_j \omega_k \), substituting (146) and then (147), assuming no torsion, show that the components of the Riemann tensor are given by

\[
R_{ijkl}^l = \Gamma^l_{ik,j} + \Gamma^m_{ik} \Gamma^l_{mj} - (i \leftrightarrow j).
\]

10. By writing \( X = X^a e_a, \text{ etc.} \), proved the Ricci identity (186) by calculation directly from the derivative of the curvature tensor.

11. Show that on a 2-dimensional manifold the Riemann tensor with all indices lowered must take the form

\[
R_{abcd} = R_{g_{aj} g_{dk} b}.
\]
12. Explain briefly why on any 2-dimensional manifold, for any given metric, there exists a coordinate system in which that metric takes the form

\[ g_{ij} = \Omega^2(x)\delta_{ij}. \]  

(221)

Calculate the Ricci scalar \( R \) in this coordinate system.

13. Prove the first and second Bianchi identities by working in normal coordinates.
4 Differential forms

4.1 Exterior algebra

A totally antisymmetric covariant $(p, 0)$ tensor field $\alpha$ is called a $p$-form. In abstract index notation we write $\alpha$ as $\alpha_{a_1 \ldots a_p}$, with the property that if $\pi$ is a permutation of $p$ objects then

$$\alpha_{\pi(a_1 \ldots a_p)} = (\text{sign } \pi)\alpha_{a_1 \ldots a_p},$$

reflecting the fact that $\alpha$ is totally antisymmetric.

We denote the set of $p$-forms by $\Lambda^p(M)$ (this is a module over $C^\infty(M)$) and the set of $p$-forms at the point $x$ by $\Lambda^p_x(M)$ (which forms a vector space over $\mathbb{R}$).

Note: By convention we take $\Lambda^0(M)$ to be the space of (smooth) functions, so that a 0-form is just a scalar field on $M$.

Example: A 2-form in $\mathbb{R}^3$ can be written out as

$$\alpha = \alpha_{12} dx \otimes dy + \alpha_{13} dx \otimes dz + \alpha_{21} dy \otimes dx + \alpha_{23} dy \otimes dz + \alpha_{31} dz \otimes dx + \alpha_{32} dz \otimes dy,$$

where $\alpha_{12} = -\alpha_{21}$, $\alpha_{13} = -\alpha_{31}$ and $\alpha_{23} = -\alpha_{32}$. We call $\alpha_{ij}$ the components of the 2-form.

Definition: Given $\alpha \in \Lambda^p(M)$ and $\beta \in \Lambda^q(M)$ we define the exterior product $\alpha \wedge \beta \in \Lambda^{p+q}(M)$ by

$$(\alpha \wedge \beta)(V_1, \ldots, V_{p+q}) = \frac{1}{p!q!} \sum_\pi (\text{sign } \pi) (\alpha(V_1, \ldots, V_p)\beta(V_{p+1}, \ldots, V_{p+q})).$$

(224)

Warning: Some books use a convention in which one divides by $(p+q)!$ rather than $p!q!$. Books that use this convention have different factors of $p, q$ and $n$ for most of the formulae in this chapter. The convention we use is that of “Analysis, Manifolds and Physics” which will be my main reference for this section. It is also the convention used in “Geometrical Methods of Mathematical Physics” by Schutz.

We can also write the above formula for the exterior product of a $p$-form $\alpha_{a_1 \ldots a_p}$ and a $p$-form $\beta_{a_1 \ldots a_q}$ in terms of abstract indices as:

$$(\alpha \wedge \beta)_{a_1 \ldots a_{p+q}} = \frac{(p+q)!}{p!q!} \alpha_{[a_1 \ldots a_p} \beta_{a_{p+1} \ldots a_{p+q}]}.$$  

(225)

Note: The numerical factor is just 1 with the alternative convention.

Note: We can use the antisymmetry of the components to write

$$\alpha = 2\alpha_{12} dx \wedge dy + 2\alpha_{13} dx \wedge dz + 2\alpha_{23} dy \wedge dz,$$

(226)

where we now only have $\alpha_{ij}$ with $i < j$. $2\alpha_{12}$, $2\alpha_{13}$ and $2\alpha_{13}$ are called the strict components of $\alpha$. (See “Analysis, Manifolds and Physics”, p. 196 for more details on this).

In the special case where $\alpha$ and $\beta$ are 1-forms the above formula (224) gives

$$\alpha \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha.$$  

(227)
Lemma: (Properties of the exterior product)
\[\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha,\]
\[\alpha \wedge \beta \wedge \gamma = \alpha \wedge (\beta \wedge \gamma),\]
\[\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma.\]

The set of forms (of all degrees) together with the exterior product is a graded algebra called the Grassman algebra and denoted \(\Lambda(M)\).

Let \(\{\theta^a(x)\}_{a=1}^n\) be a basis for \(T^*_x(M)\) and let \(\alpha \in \Lambda^p(M)\). Then we have
\[\alpha = \alpha_{a_1 \ldots a_p} \theta^{a_1} \otimes \cdots \otimes \theta^{a_p}\]
\[= \sum_{a_1 < a_2 < \cdots < a_p} \alpha_{a_1 \ldots a_p} \theta^{a_1} \wedge \cdots \wedge \theta^{a_p}.\]

We see that \(\{\theta^{a_1} \wedge \cdots \wedge \theta^{a_p} : a_1 < a_2 < \cdots < a_p\}\) is a basis for \(\Lambda^p(M)\). Hence
\[\dim \Lambda^p(M) = \frac{n!}{(n-p)!p!}.\]

Remark: In calculations it is often useful to write
\[\alpha = \frac{1}{p!} \alpha_{a_1 \ldots a_p} \theta^{a_1} \wedge \cdots \wedge \theta^{a_p}.\]

This is a perfectly good expression for \(\alpha\), but without insisting that the indices are ordered by requiring \(a_1 < a_2 < \cdots < a_p\) the set \(\{\theta^{a_1} \wedge \cdots \wedge \theta^{a_p}\}\) is not a basis for \(\Lambda^p(M)\) since, although it spans \(\Lambda^p(M)\), the various elements are not independent. For example \(\theta^{a_1} \wedge \theta^{a_2} \cdots \wedge \theta^{a_p} = -\theta^{a_2} \wedge \theta^{a_1} \cdots \wedge \theta^{a_p}\).

4.2 The exterior derivative

In Chapter 2 we have seen how to take the directional derivative \(X(f)\) of a function \(f\) in the direction \(X\) (in a coordinate basis it is just given by \(X(f) = X^i \frac{\partial f}{\partial x^i}\)).

We may use this to define the exterior derivative of the function \(f\) to be the 1-form \(df\) given by
\[df(X) := X(f).\]

Since in a coordinate basis we have \(X(f) = X^i \frac{\partial f}{\partial x^i}\), then in a coordinate basis the exterior derivative of a function must be given by
\[df = \frac{\partial f}{\partial x^i} dx^i.\]

We want to extend \(d\) to a derivative operator which takes \(p\)-forms to \((p+1)\)-forms which has the following properties:

1. \(d(\alpha + \beta) = d\alpha + d\beta\) and for constant \(c\) \(d(c\alpha) = cd(\alpha)\); i.e. \(d\) is \(\mathbb{R}\)-linear;
2. \(d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta\); i.e. \(d\) satisfies a Leibniz type rule;
3. \(d^2 = 0\);
4. If \(f\) is a 0-form then \(df\) is the usual derivative (as given above).
5. The operation is local so that if \( U \) is an open set and \( \alpha \) and \( \beta \) are differential forms that agree on \( U \) then so do their exterior derivatives.

We now show that if such an operator \( d \) exists, then 1.-4. imply that the local expression for \( d\alpha \) is uniquely defined (and in the course of the proof we will derive the unique local expression). One can then show that the local expression we have obtained does indeed satisfy 1.-5. and also transforms correctly under a change of coordinates. Finally by making use of 5. one can extend this local formula to a global definition on the manifold. An alternative approach is to simply give a coordinate free global definition for the exterior derivative but this is not very illuminating at first sight (see the formula for \( d \) in Section 4.3 involving the Lie bracket).

We start with the proof that properties 1.-4. define \( d \) uniquely in a given coordinate system. Let \( \alpha \in \Lambda^p(M) \), then in terms of local coordinates \( x^i \) we can write it in terms of its components as

\[
\alpha = \alpha_{i_1...i_p} \, dx^{i_1} \otimes \cdots \otimes dx^{i_p} = \frac{1}{p!} \alpha_{i_1...i_p} \, dx^{i_1} \wedge \cdots \wedge dx^{i_p}.
\]  

(237)

Hence by properties 1., 2. and 3.,

\[
d\alpha = \frac{1}{p!} \left( d\alpha_{i_1...i_p} \right) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}.
\]  

(238)

But the coefficients \( \alpha_{i_1...i_p} \) are just functions so by property 5. we must have

\[
d\alpha = \frac{1}{p!} \left( \partial \alpha_{i_1...i_p} / \partial x^j \right) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} = \frac{(p+1)!}{p!} \alpha_{[i_1...i_p,j]} \, dx^j \otimes dx^{i_1} \cdots \otimes dx^{i_p}.
\]  

(239)

Hence the local coordinate expression for \( d\alpha \) must uniquely be given by

\[
(d\alpha)_{i_1...i_{p+1}} = (p+1)\alpha_{[i_2...i_{p+1},i_1]}.
\]  

(240)

Note the derivative index comes first! (Note also that there is no factor of \((p+1)\) with the alternative convention.)

Using the above formula, it is not hard to check properties 1.-5. The definition is clearly local so that 5. is also satisfied. After a rather long calculation one can also check that the formula above transforms correctly under a change of coordinates - the key point is that the terms involving the second derivatives of the coordinate transformation \( \partial^2 x^i / \partial x^j \partial x^k \), which prevent the ordinary partial derivative of a tensor transforming as a tensor vanish as a result of the antisymmetrisation in (240).

We may therefore define the exterior derivative in any chart using formula (240), but on the overlap they must agree by the local uniqueness and the fact that the local expression transforms correctly. Hence we can piece the local expression together to give a global formula.

**Definition:** A differential \( p \)-form \( \alpha \) is **exact** if there exists a \((p-1)\)-form \( \beta \) such that

\[
\alpha = d\beta.
\]  

(241)
Clearly if $\alpha$ and $\tilde{\alpha}$ are exact then for constants $c_1, c_2 \in \mathbb{R}$ then $c_1\alpha + c_2\tilde{\alpha}$ is also exact. The vector space of exact $p$-forms is denoted $B^p(M)$.

**Definition:** A $p$-form $\alpha$ is **closed** if

$$d\alpha = 0.$$  \hfill (242)

The space of closed $p$-forms again forms a vector space (over $\mathbb{R}$) which is denoted $Z^p(M)$.

Because $d^2 = 0$ it follows that every exact form is closed (so that $B^p(M)$ forms a vector subspace of $Z^p(M)$). It turns out that the converse is not in general true - so that not all closed forms need be exact. This will be the topic of Section 6 where we study De Rham cohomology. There, the quotient spaces

$$H^p(M) = Z^p(M)/B^p(M), \quad p = 0, \ldots, n.$$  \hfill (243)

will give us topological information about the manifold $M$.

**Examples:**

1. Let $f$ be a 0-form on $\mathbb{R}^3$. Then

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$  \hfill (244)

So that we see the components of $df$ are just the same as the components of the classical grad $f$ in 3-dimensions. Note however $df$ (and indeed grad $f$) transforms as a 1-form (or co-vector), not as a vector field. There is no difference in Euclidean coordinates but there is a difference if one uses general curvilinear coordinates.

2. Let $\alpha$ be the 1-form on $\mathbb{R}^3$ given in components by

$$\alpha = A(x,y,z)dx + B(x,y,z)dy + C(x,y,z)dz.$$  \hfill (245)

Then

$$d\alpha = \frac{\partial A}{\partial y} dy \wedge dx + \frac{\partial A}{\partial z} dz \wedge dx + \frac{\partial B}{\partial x} dx \wedge dy$$
$$+ \frac{\partial B}{\partial z} dz \wedge dy + \frac{\partial C}{\partial x} dx \wedge dz + \frac{\partial C}{\partial y} dy \wedge dz$$
$$= \left( \frac{\partial C}{\partial y} - \frac{\partial B}{\partial z} \right) dy \wedge dz + \left( \frac{\partial A}{\partial z} - \frac{\partial C}{\partial x} \right) dz \wedge dx$$
$$+ \left( \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx \wedge dy.$$  \hfill (246)

If we think of $(A, B, C)$ as the components of a vector $v$ then the components of $d\alpha$ are the same as the components of curl $v$ (in Euclidean space $\mathbb{R}^3$ only).

3. Let $\beta$ be a 2-form on $\mathbb{R}^3$ with components

$$\beta = P(x,y,z)dy \wedge dz + Q(x,y,z)dz \wedge dx + R(x,y,z)dx \wedge dy.$$  \hfill (247)
Then
\[ d\beta = \frac{\partial P}{\partial x} \, dx \wedge dy \wedge dz + \frac{\partial Q}{\partial y} \, dy \wedge dz \wedge dx + \frac{\partial R}{\partial z} \, dz \wedge dx \wedge dy \]
\[ = \left( \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \wedge dy \wedge dz. \] (248)

If we think of \((P, Q, R)\) as the components of a vector field \(v\) then \(d\beta\) is the same as \(\text{div} \, v\) (in Euclidean space \(\mathbb{R}^3\) only).

4. Using the fact that \(d^2 = 0\) and the above identifications of forms with scalar and vector fields we see that
\[ \text{curl}(\text{grad} \, f) = 0, \quad \text{div}(\text{curl} \, v) = 0. \] (249)

### 4.3 The pull-back of differential forms

In Section 2.2 we saw how to use a map
\[ f : M \to N \] (250)
to pull back a form \(\omega \in \Lambda^p(N)\) to a form \(f^*\omega \in \Lambda^p(M)\) by defining
\[ (f^*\omega)(V_1, \ldots, V_p) = \omega(f_*V_1, \ldots, f_*V_p). \] (251)

In components if \(x^i\) are local coordinates on \(M\) and \(y^\alpha = f(x^1, \ldots, x^m)\) then
\[ (f^*\omega)_{i_1 \ldots i_p} = \frac{\partial y^{\alpha_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{\alpha_p}}{\partial x^{i_p}} \omega_{\alpha_1 \ldots \alpha_p}. \] (252)

**Theorem:**
\[ f^*(\alpha + \beta) = f^*(\alpha) + f^*(\beta), \]
\[ f^*(\alpha \wedge \beta) = f^*(\alpha) \wedge f^*(\beta), \]
\[ f^*(d\alpha) = d(f^*\alpha). \] (253) (254) (255)

**Proof:** The proof of 1. and 2. is a straightforward calculation. Since we may write any \(p\)-form as a sum of products of 0-forms with exact 1-forms, 1. and 2. mean that we need only prove 3. for such forms.

If \(\alpha\) is a 0-form \(\phi\) then
\[ d(f^*\phi) = d(\phi \circ f) = d\phi \circ Df = f^*d\phi \] (256)

If \(\alpha\) is an exact 1-form \(\alpha = d\phi\) then
\[ f^*(d\alpha) = f^*(d^2\phi) = 0. \] (257)

On the other hand
\[ d(f^*\alpha) = d(f^*d\phi) = d(d(f^*\phi) = d^2(f^*\phi) = 0. \] (258)
4.4 Interior product

Let \( \omega \in \Lambda^p(M) \) be a \( p \)-form, and let \( X \) be a vector field. Then we may define a \((p - 1)\)-form \( \iota_X \omega \), called the **interior product** of \( \omega \) with \( X \), by

\[
\iota_X \omega (V_1, \ldots, V_{p-1}) = \omega(X, V_1, \ldots, V_{p-1}).
\]

In components, if

\[
\omega = \omega_{i_1 \ldots i_p} dx^{i_1} \otimes \cdots \otimes dx^{i_p},
\]

then

\[
\iota_X \omega = X^j \omega_{ji_2 \ldots i_p} dx^{i_2} \otimes \cdots \otimes dx^{i_p} = \frac{1}{(p-1)!} X^j \omega_{ji_2 \ldots i_p} dx^{i_2} \wedge \cdots \wedge dx^{i_p}.
\]

The interior product is simply the contraction with \( X \). In the case of a 0-form \( \phi \) we define \( \iota_X \phi = 0 \).

**Proposition:** Let \( \alpha \in \Lambda^p(M) \) and \( \beta \in \Lambda^q(M) \). Then

\[
\iota_X (\alpha + \beta) = \iota_X \alpha + \iota_X \beta,
\]

\[
\iota_X (\alpha \wedge \beta) = (\iota_X \alpha \wedge \beta + (-1)^p \alpha \wedge \iota_X \beta),
\]

\[
\iota_X^2 = 0.
\]

**Proof:** The proof of 1. and 2. is a straightforward calculation. The proof of 3. comes from the contraction of \( X^i X^j \) (which is symmetric) with the antisymmetric components of the form.

4.5 Lie derivative of differential forms

We already know that for tensor fields \( S \) and \( T \) we have

\[
\mathcal{L}_V (S \otimes T) = \mathcal{L}_V S \otimes T + S \otimes \mathcal{L}_V T.
\]

If we take \( S \) and \( T \) to be differential forms \( \alpha \) and \( \beta \) and we antisymmetrise the resulting equation we obtain

\[
\mathcal{L}_V (\alpha \wedge \beta) = \mathcal{L}_V \alpha \wedge \beta + \alpha \wedge \mathcal{L}_V \beta.
\]

We also know from (255) that the exterior derivative commutes with the pull back. In particular we may take \( f^* \) to be the (local) 1-parameter group of diffeomorphisms generated by a vector field \( V \), so that

\[
d(f^* \alpha) = f^* (d \alpha).
\]

Differentiating this expression with respect to \( t \) and setting \( t = 0 \) gives

\[
d(\mathcal{L}_V \alpha) = \mathcal{L}_V (d \alpha),
\]

and hence \( d \) also commutes with the Lie derivative.
**Proposition:** (Cartan’s formula for the Lie derivative in terms of the exterior derivative)
\[ \mathcal{L}_V = d \circ \iota_V + \iota_V \circ d \]  
(269)

**Proof:** We know that \( \mathcal{L}_V (\alpha \wedge \beta) = \mathcal{L}_V \alpha \wedge \beta + \alpha \wedge \mathcal{L}_V \beta \). On the other hand
\[
(d \circ \iota_V + \iota_V \circ d)(\alpha \wedge \beta) = d [\iota_V (\alpha) \wedge \beta + (-1)^p \alpha \wedge \iota_V (\beta)] + \iota_V [d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta]
\]
\[
= [(d \circ \iota_V + \iota_V \circ d)\alpha] \wedge \beta + (-1)^p (\iota_V \alpha) \wedge (d\beta) + (-1)^p (d\alpha) \wedge (\iota_V \beta)
\]
\[
+ (-1)^{p+1} (d\alpha) \wedge (\iota_V \beta) + (-1)^p (\iota_V \alpha) \wedge (d\beta) + (-1)^{2p} \alpha \wedge [(d \circ \iota_V + \iota_V \circ d)\beta]
\]
\[
= [(d \circ \iota_V + \iota_V \circ d)\alpha] \wedge \beta + \alpha \wedge [(d \circ \iota_V + \iota_V \circ d)\beta].
\]  
(270)

Hence both \( \mathcal{L}_V \) and \( (d \circ \iota_V + \iota_V \circ d) \) respect the exterior product. Since any form is a sum of products of 0-forms and exact 1-forms and both the operators respect the exterior product and are linear we only need to verify (269) for 0-forms and exact 1-forms.

1. \( \mathcal{L}_V f = V(f) = \iota_V df = (\iota_V \circ d + d \circ \iota_V) f \), since \( \iota_V f = 0 \) by definition.
2. \( \mathcal{L}_V df = d\mathcal{L}_V f = d(V(f)) = d(\iota_V df) = (\iota_V \circ d + d \circ \iota_V) df \), since \( d^2 f = 0 \).

**Proposition:** (Alternative, global, definition of the exterior derivative) Let \( \omega \) be a \( p \)-form. Then
\[
d\omega(V_0, V_1, \ldots, V_p) = \sum_{\alpha=0}^p (-1)^\alpha V_\alpha \left( \omega(V_0, \ldots, \hat{V}_\alpha, \ldots, V_p) \right) + \sum_{0 \leq \alpha < \beta \leq p} (-1)^{\alpha + \beta} \omega([V_\alpha, V_\beta], V_0, \ldots, \hat{V}_\alpha, \hat{V}_\beta, \ldots, V_p),
\]  
(271)

where the hat on \( \hat{V}_\alpha \) indicates that these terms are omitted. (Note the \( V_\alpha \) have to be vector fields, rather than vectors at a point, for each term in this definition to be defined. However, the sum of all terms contains only the \( V_\alpha \) at a point.)

**Proof:** Calculate in local coordinates! Rather than do this in general we give the above formula in the special case of a 1-form and show that this agrees with the standard result in local coordinates. For a 1-form \( \theta \) the above formula reduces to
\[
d\theta(V, W) = \mathcal{L}_V(\theta(W)) - \mathcal{L}_W(\theta(V)) - \theta([V, W]).
\]  
(272)

In local coordinates we write \( \theta = \theta_i dx^i \), \( V = V^i \partial / \partial x^i \) and \( W = W^i \partial / \partial x^i \). Then \( d\theta = \theta_{i,j} dx^j \wedge dx^i + \theta_{i,j} (dx^j \otimes dx^i - dx^i \otimes dx^j) \). Hence
\[
d\theta(V, W) = (\alpha_{i,j} - \alpha_{j,i}) W^i V^j.
\]  
(273)

On the other hand
\[
\mathcal{L}_V(\theta(W)) = \mathcal{L}_V(\theta_i W^i) = V^j \theta_{i,j} W^i + \theta_i V^j W^i_{,j},
\]  
(274)

\[
\mathcal{L}_W(\theta(V)) = \mathcal{L}_W(\theta_i V^i) = W^j \theta_{i,j} V^i + \theta_i W^j V^i_{,j},
\]  
(275)

\[
\theta([V, W]) = \theta_i [V, W]^i = \theta_i V^j W^i_{,j} - \theta_i W^j V^i_{,j}.
\]  
(276)

Hence
\[
\mathcal{L}_V(\theta(W)) - \mathcal{L}_W(\theta(V)) - \theta([V, W]) = (\alpha_{i,j} - \alpha_{j,i}) W^i V^j = d\theta(V, W).
\]  
(277)
4.6 The Cartan structural equations

In this section, we work with an arbitrary basis of vectors $e_a$ and the dual basis of 1-forms $\theta^a$, such that $\theta^a_b e^b = \delta^a_b$.

Recall the definition (141) of the connection coefficients, which is equivalent to

$$\Gamma^a_{bc} = \theta^a(\nabla_b e_c)$$  \hspace{1cm} (278)

(Warning: in some books the $b$ and $c$ indices are written the other way round), and the definition (50) of the structure constants, which is equivalent to

$$C^a_{bc} = \theta^a([e_b, e_c]).$$  \hspace{1cm} (279)

Recall also the definition (160) of the torsion tensor and the relation (162) between the torsion, connection coefficients and structure constants. Finally, recall the definition (169) of the curvature tensor, and the expression (182) for the curvature tensor in terms of the connection coefficients and their derivatives in a general (non-coordinate) basis.

Since $T^a_{bc}$ is antisymmetric in $b$ and $c$ we can define the torsion 2-forms $\Theta^a$ (for $a = 1, \ldots, n$) by

$$\Theta^a := \frac{1}{2} T^a_{bc} \theta^b \wedge \theta^c,$$  \hspace{1cm} (280)

and since $R^a_{bcd}$ is antisymmetric in $c$ and $d$ we can define the curvature 2-forms

$$\Omega^a_{bc} := \frac{1}{2} R^a_{bcd} \theta^c \wedge \theta^d.$$  \hspace{1cm} (281)

We also define the connection 1-forms $\omega^a_b$ by

$$\omega^a_b := \Gamma^a_{cb} \theta^c.$$  \hspace{1cm} (282)

(Note that using our convention for the definition of $\Gamma^a_{bc}$ the contraction is on the first index).

**Remark:** In this approach the indices on $\Theta^a$, $\Omega^a_{bc}$ and $\omega^a_b$ are not tensor indices but simply labels. However it is instructive to see what happens to these under a change of basis to

$$e_{a'} = L^b_{a'} e_b, \quad \theta^{a'} = \tilde{L}^b_{a'} \theta^b, \quad \text{where } \tilde{L} = L^{-1}$$  \hspace{1cm} (283)

Since $T^a_{bc}$ are the components of a tensor,

$$\Theta^{a'} = \tilde{L}^a_{a'} \theta^b.$$  \hspace{1cm} (284)

Similarly since $R^a_{bcd}$ are the components of a tensor

$$\Omega^{a'}_{b'd} = \tilde{L}^a_{a'} c^d_{b'} \Omega^c_d.$$  \hspace{1cm} (285)

However $\Gamma^a_{bc}$ are **not** the components of a tensor. This means that $\omega^a_b$ transforms as (exercise)

$$\omega^{a'}_{b'} = \tilde{L}^a_{a'} c^d_{b'} \omega^c_d + d\tilde{L}^a_{a'} c^d_{b'}. $$  \hspace{1cm} (286)

The $d\tilde{L}^a_{a'} c^d_{b'} \omega^c_d$ term being the usual additional inhomogeneous term one gets from the components of a connection under a change of basis.
**Theorem:** (Cartan structural equations)

\[
\Theta^a = d\theta^a + \omega^a_b \wedge \theta^b, \quad (287)
\]

\[
\Omega^a_b = d\omega^a_b + \omega^a_m \wedge \omega^m_b \quad (288)
\]

(The right-hand sides of these equations are sometimes written as \(D\theta^a\) and \(D\omega^a_b\) respectively, in self-explanatory notation.)

**Proof:** To prove (287) we first note that from the result of the previous section

\[
d\theta(V, W) = L_V(\theta(W)) - L_W(\theta(V)) - \theta([V, W]). \quad (289)
\]

Taking \(\theta = \theta^a\), \(V = e_b\) and \(W = e_c\) gives

\[
d\theta^a(e_b, e_c) = e_b(\delta^a_c) - e_c(\delta^a_b) - C^a_{bc} = -C^a_{bc}, \quad (290)
\]

so that

\[
d\theta^a = -\frac{1}{2} C^a_{bc} \theta^b \wedge \theta^c. \quad (291)
\]

Also,

\[
\omega^a_b \wedge \theta^b = \Gamma^a_{cb} \theta^e \wedge \theta^b = \frac{1}{2} (\Gamma^a_{bc} - \Gamma^a_{cb}) \theta^b \wedge \theta^c. \quad (292)
\]

Hence

\[
d\omega^a_b + \omega^a_b \wedge \theta^b = \frac{1}{2} (\Gamma^a_{cb} - \Gamma^a_{cb} - C^a_{bc}) \theta^b \wedge \theta^c = \frac{1}{2} T^a_{bc} \theta^b \wedge \theta^c \quad (293)
\]

which proves the first of the Cartan structural equations.

To prove (288) we first note that \(\omega^a_b = \Gamma^a_{cb} \theta^e\), so that

\[
d\omega^a_b = \Gamma^a_{cb} \omega^e_b \wedge \theta^e + \Gamma^a_{cb} d\theta^e. \quad (294)
\]

Now

\[
(e_d \Gamma^a_{cb}) \theta^d = e_d \Gamma^a_{cb} \theta^j dx^j = \delta^d_j \Gamma^a_{cb} \theta^j dx^j = \Gamma^a_{cb} dx^j, \quad (295)
\]

where the second equality holds since \(e^i_d\) and \(\theta^j_d\) are inverses of each other. We also have

\[
d\theta^e = -\frac{1}{2} C^e_{de} \theta^d \wedge \theta^e. \quad (296)
\]

Hence

\[
d\omega^a_b = e_d (\Gamma^a_{cb}) \theta^d \wedge \theta^e - \frac{1}{2} \Gamma^a_{cb} C^e_{de} \theta^d \wedge \theta^e
\]

\[
= \frac{1}{2} [e_c (\Gamma^a_{db}) - e_d (\Gamma^a_{cb}) - \Gamma^a_{cb} C^e_{de}] \theta^c \wedge \theta^d. \quad (297)
\]

Finally

\[
\omega^a_e \wedge \omega^e_b = (\Gamma^a_{ce} \theta^e) \wedge (\Gamma^e_{db} \theta^d)
\]

\[
= \Gamma^a_{ce} \Gamma^e_{db} \theta^e \wedge \theta^d
\]

\[
= \frac{1}{2} (\Gamma^a_{ce} \Gamma^e_{db} - \Gamma^a_{db} \Gamma^e_{ce}) \theta^e \wedge \theta^d. \quad (298)
\]
Hence
\[ \begin{align*}
   d\omega^a_b + \omega^a_c \wedge \omega^c_b &= 1  \\
   &= \frac{1}{2} [e_c (\Gamma^a_{db}) - e_d (\Gamma^a_{cb}) + \Gamma^a_{ce} \Gamma^e_{db} - \Gamma^a_{de} \Gamma^e_{cb} - C^e_{cd} \Gamma^a_{eb} \theta^c \wedge \theta^d]  \\
   &= \frac{1}{2} R^a_{bcd} \theta^c \wedge \theta^d  \\
   &= \Omega^a_b,  \\
\end{align*} \] (299)

which proves the second of the Cartan structural equations.

**Remark:** From the Cartan viewpoint Eqs. (287) and (288) are the natural definitions of the torsion 2-form \( \Theta^a_b \) and the curvature 2-form \( \Omega^a_{bc} \). The hard part in proving the above theorem is relating the Cartan approach to the Koszul approach where the connection is given in terms of \( \nabla_U V \).

Differentiating the first structural equation gives
\[ \begin{align*}
   d\Theta^a &= d\omega^a_b \wedge \theta^b - \omega^a_c \wedge d\theta^b  \\
   &= \Omega^a_b \wedge \theta^b - \omega^a_c \wedge \omega^c_b \wedge \theta^b - \omega^a_b \wedge \Theta^b + \wedge^a_b \wedge \omega^c_b \wedge \theta^c  \\
   &= \Omega^a_b \wedge \theta^b - \omega^a_b \wedge \Theta^b.  \\
\end{align*} \] (300)

**Corollary:** (First Bianchi identity) For a torsion-free connection (300) gives
\[ \Omega^a_b \wedge \theta^b = 0. \] (301)

In terms of \( R^a_{bcd} \), this is just
\[ R^a_{bcd} \wedge \theta^c \wedge \theta^d = 0. \] (302)

Hence \( R^a_{[bcd]} = 0 \) or using the fact that \( R^a_{bcd} \) is antisymmetric on the last 2 indices
\[ R^a_{bcd} + R^a_{cde} + R^a_{edc} = 0, \] (303)

which is the first Bianchi identity.

**Proposition:** (Second Bianchi identity) Differentiating the second Cartan equation gives
\[ d\Omega^a_b = \Omega^a_{c} \wedge \omega^c_b - \omega^a_c \wedge \Omega^c_b. \] (304)

**Proof:**
\[ \begin{align*}
   d\Omega^a_b &= d\omega^a_c \wedge \omega^c_b - \omega^a_d \wedge d\omega^c_b  \\
   &= \Omega^a_c \wedge \omega^c_b - \omega^a_d \wedge \omega^d_m \wedge \omega^e_b - \omega^a_e \wedge \Omega^e_b + \omega^a_c \wedge \omega^c_d \wedge \omega^d_b  \\
   &= \Omega^a_c \wedge \omega^c_b - \omega^a_c \wedge \Omega^c_b.  \\
\end{align*} \] (305)

**Remark:** In terms of \( R^a_{bcd} \) in the torsion-free case this is just the second Bianchi identity (exercise)
\[ R^a_{bcd,e} + R^a_{bde,c} + R^a_{bec,d} = 0. \] (306)
4.7 Exercises

1. Show that in the definition
   \[ d\alpha = \frac{1}{p!} \partial_j \alpha_{i_1 \ldots i_p} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \]  
   (307)

   of the exterior derivative \( d \) of differential forms, the partial derivative \( \partial_i \) can be replaced by a covariant derivative \( \nabla_i \) (related to \( \partial_i \) by the connection coefficients \( \Gamma^i_{jk} \)), under one further assumption that you must state.

2. From suitable definitions and results in the lecture notes, derive the transformation rule
   \[ \omega^a_{\nu'} = \tilde{L}^a_{\nu'} \tilde{L}^d_{\nu'} \omega^d + d\tilde{L}^a_{\nu'} \tilde{L}^d_{\nu'} . \]  
   (308)

   of the connection 1-form \( \omega^a_b \) under a change of (general) basis.

3. From suitable definitions and results in the lecture notes, derive the second Bianchi identity for torsion-free connections in its abstract index form
   \[ R^a_{bcd,e} + R^a_{bde,c} + R^a_{bec,d} = 0 \]  
   (309)

   from its differential forms form
   \[ d\Omega^a_b = \Omega^a_c \wedge \omega^c_b - \omega^a_c \wedge \Omega^c_b . \]  
   (310)

4. Prove that with the time-dependent vector field \( V_t \) defined as the time derivative of the flow \( \phi_t \), that is \( d\phi_t / dt = V_t(\phi_t(x)) \), the identity
   \[ \frac{d}{dt} \phi_t^* \alpha = \phi_t^* \mathcal{L}_{V_t} \alpha \]  
   (311)

   holds for any differential form \( \alpha \).
5 Integration of differential forms

In this section we will be considering the integration of an \( n \)-form \( \omega \) over an \( n \)-dimensional paracompact manifold \( M \), written as

\[ \int_M \omega. \]  

(312)

For reasons that will be clear later we will only consider integration over orientable manifolds \( M \). It was shown by de Rham that is possible to extend the theory of differential forms to allow integration over non-orientable manifolds but we will not consider this here.

Two coordinate systems \( x^i \) and \( \tilde{x}^i \) on an opens set \( U \) of \( \mathbb{R}^n \) are said to have the same orientation if the Jacobian determinant is positive, so that

\[ \tilde{J}(x) = \det \left( \frac{\partial x^i}{\partial \tilde{x}^j} \right) > 0 \quad \forall x \in U. \]  

(313)

**Definition:** A manifold \( M \) is **orientable** if there exists an atlas such that on the overlap between two charts \((U, \varphi)\) and \((V, \varphi')\) the Jacobian determinant between the local coordinates is positive everywhere on \( U \cap V \).

5.1 The alternating symbol and the determinant

Before starting on the theory of integration of \( n \)-forms it is useful to introduce a new piece of notation.

**Definition:** The alternating symbol in \( n \) dimensions \( \epsilon_{i_1 \ldots i_n} \) is defined by

\[ \epsilon_{i_1 \ldots i_n} = \begin{cases} +1 & \text{if } i_1, \ldots, i_n \text{ is an even permutation of } 1, \ldots, n, \\ -1 & \text{if } i_1, \ldots, i_n \text{ is an odd permutation of } 1, \ldots, n, \\ 0 & \text{otherwise}. \end{cases} \]  

(314)

Note that unlike the Kronecker delta \( \delta^i_j \) the alternating symbol does not correspond to the components of a tensor in an arbitrary coordinate system.

**Definition:** Let \( A = \{ A^i_j \} \) be an \( n \times n \) matrix. Then we may define the determinant of \( A \) according to

\[ \det A = \epsilon_{i_1 \ldots i_n} A^{i_1} A^{i_2} \ldots A^{i_n}. \]  

(315)

One can use induction on \( n \) to show that this is equivalent to the recursive definition of the determinant in terms of a co-factor expansion.

We now use the above result to calculate how an \( n \)-form transforms under a change of local coordinates. Let \( \omega \) be an \( n \)-form and \( x^i \) a local coordinate system. We may write \( \omega \) in terms of its components in this coordinate system as

\[ \omega = \omega_{i_1 \ldots i_n} dx^{i_1} \otimes \cdots \otimes dx^{i_n}. \]  

(316)

Since \( \omega_{i_1 \ldots i_n} \) is totally antisymmetric we can write

\[ \omega_{i_1 \ldots i_n} = \omega^{i_1 \ldots n} \epsilon_{i_1 \ldots i_n}. \]  

(317)
If we now make a change to a new local coordinate system $\tilde{x}^i$ then according to the usual formula for the change of coordinates of a tensor we have

$$\tilde{\omega}_{1...n} = \frac{\partial x^i}{\partial \tilde{x}^1} \frac{\partial x^j}{\partial \tilde{x}^2} \cdots \frac{\partial x^n}{\partial \tilde{x}^n} \omega_{i_1...i_n} = \frac{\partial x^i}{\partial \tilde{x}^1} \frac{\partial x^j}{\partial \tilde{x}^2} \cdots \frac{\partial x^n}{\partial \tilde{x}^n} \epsilon_{i_1...i_n} \omega_{1...n} = \det \left( \frac{\partial x^i}{\partial \tilde{x}^j} \right) \omega_{1...n}. \quad (318)$$

Hence we have shown that

$$\tilde{\omega}_{1...i_n} = \det \left( \frac{\partial x^i}{\partial \tilde{x}^j} \right) \omega_{1...i_n}. \quad (319)$$

### 5.2 The definition of the integral

The support of a function is the closure of the set on which it is non-zero. We start by considering the definition of the integral of an $n$-form which has support contained in the domain of a single chart $U$ with local coordinates $x^i$. In this situation we define

$$\int_M \omega = \int_{\mathbb{R}^n} \omega_{1...n}(x^1, ..., x^n) d^n x, \quad (320)$$

where the right hand integral is just the Lebesgue integral of the function $\omega_{1...n}(x^1, ..., x^n)$ over $\mathbb{R}^n$.

Suppose now that we introduce a new set of local coordinates $\tilde{x}^i$ with the same orientation on the set $U$. Then

$$\int_{\mathbb{R}^n} \tilde{\omega}_{1...n}(\tilde{x}^1, ..., \tilde{x}^n) d^n \tilde{x} = \int_{\mathbb{R}^n} \omega_{1...n}(x^1(\tilde{x}^k), ..., x^n(\tilde{x}^k)) \det \left( \frac{\partial x^i}{\partial \tilde{x}^j} \right) d^n \tilde{x}$$

$$= \int_{\mathbb{R}^n} \omega_{1...n}(x^1, ..., x^n) d^n x$$

$$= \int_M \omega, \quad (321)$$

so that the value of the integral does not depend upon the choice of local coordinates. Note that if $\tilde{x}^i$ did not have the same orientation there would be an overall change of sign.

We now define the integral of an $n$-form on $M$ using a partition of unity to piece together the local integrals. Let $\mathcal{A}$ be an atlas for $M$ and let $\phi_\lambda$ be a partition of unity subordinate to the atlas. We now demand that the integral is linear so that

$$\int_M (c_1 \omega + c_2 \tilde{\omega}) = c_1 \int_M \omega + c_2 \int_M \tilde{\omega}, \quad c_1, c_2 \in \mathbb{R}. \quad (322)$$

We may now use the partition of unity to write

$$\omega = \sum_{\lambda \in \Lambda} \phi_\lambda \omega = \sum_{\lambda \in \Lambda} (\phi_\lambda \omega). \quad (323)$$
Now for any fixed $\lambda_0$ in the above sum there exists the a chart $(U, \varphi)$ with local coordinates $x^i$ such that $\text{supp}\varphi_{\lambda_0} \omega \subset U$ and hence we may define
\[
\int_m (\varphi_{\lambda_0} \omega) = \int_{\mathbb{R}^n} (\varphi_{\lambda_0} \omega)_{1\ldots n}(x^1, \ldots, x^n) d^n x. \tag{324}
\]
Finally we use the linearity of the integral to write
\[
\int_M \omega = \int_M \left( \sum_{\lambda \in \Lambda} \varphi_\lambda \omega \right) = \sum_{\lambda \in \Lambda} \int_M (\varphi_\lambda \omega). \tag{325}
\]
We say the $n$-form is integrable if each of the $(\varphi_{\lambda_0} \omega)_{1\ldots n}(x^1, \ldots, x^n)$ is in $L^1(\mathbb{R}^n)$ (i.e. is Lebesgue-integrable) and the sum (325) converges. Note that if $M$ is compact the fact that the partition of unity is locally finite means that this sum has only a finite number of terms which shows, for example, that every continuous $n$-form $\omega$ on a compact manifold $M$ is integrable (or more generally any continuous $n$-form of compact support is integrable).

**Note:** One can show that the definition of $\int_M \omega$ does not depend on the choice of partition of unity. This is a rather straightforward (but lengthy) calculation which involves taking two partitions of unity $\varphi_\lambda$, and $\varphi_\mu$, and then looking at a partition subordinate to them both by looking at the doubly indexed partition of unity $\varphi_\lambda \varphi_\mu$.

If $S$ is some subset of $M$ (not necessarily a manifold) we define
\[
\int_S \omega = \int_M \chi_S \omega, \tag{326}
\]
where $\chi_S$ is the characteristic function of $S$ defined by
\[
\chi_S(x) = \begin{cases} 
1 & x \in S, \\
0 & \text{otherwise}. 
\end{cases} \tag{327}
\]
It follows from the definition of the integral that if two subsets $S$ and $\tilde{S}$ differ by a set of measure zero then
\[
\int_S \omega = \int_{\tilde{S}} \omega \tag{328}
\]
It also follows from the definition that the integral is additive with respect to the domain of integration,
\[
\int_{S_1 \cup S_2} \omega = \int_{S_1} \omega + \int_{S_2} \omega - \int_{S_1 \cap S_2} \omega. \tag{329}
\]

**Proposition:** Let $f : M \to N$ be an orientation preserving diffeomorphism and $\omega \in \Lambda^n(N)$ then
\[
\int_M f^* \omega = \int_{f(M)} \omega = \int_N \omega. \tag{330}
\]
Outline proof: First suppose that $\omega$ has support in a single chart of $N$ local coordinates $y^i$ and that $f^*\omega$ also has support in a single chart with local coordinates $x^j$. Then from the definition of $f^*$ the components of $f^*\omega$ are given by

$$(f^*\omega)_{1...n} = \det \left( \frac{\partial y^i}{\partial x^j} \right) \omega_{1...n},$$

and hence from the definition of the integral

$$\int_M f^*\omega = \int_{\mathbb{R}^n} (f^*\omega)_{1...n}(x^1, \ldots, x^n)d^n x$$

$$= \int_{\mathbb{R}^n} \det \left( \frac{\partial y^i}{\partial x^j} \right) \omega_{1...n}(y^1(x^k), \ldots, y^n(x^k))d^n x$$

$$= \int_{\mathbb{R}^n} \omega_{1...n}(y^1, \ldots, y^n)d^n y$$

$$= \int_N \omega.$$ (332)

One can then extend the above result from this special case to the general case by choosing a suitable atlas and partitions of unity on $M$ and $N$.

5.3 Stokes’ Theorem

In this section we prove a result which relates the derivative of an exact $n$-form $\omega = d\mu$ on a manifold with boundary to the integral of $\mu$ over the boundary of $M$.

The definition of a manifold with boundary is similar to the definition of a manifold from Section 1, but there are now two different kinds of chart. The standard chart maps $U$ onto an open set in $\mathbb{R}^n$ and a boundary chart maps an open set $U$ onto $\mathbb{R}^n+ = \{(x^1, \ldots, x^n) \in \mathbb{R}^n : x^1 \geq 0\}$. The boundary $\partial M$ is then defined as the set of points in $M$ whose images under a boundary chart lie in the boundary $\mathbb{R}^{n-1} = \{(x^1, \ldots, x^n) \in \mathbb{R}^n : x^1 = 0\}$. The set of such points form an $(n-1)$-dimensional manifold $\partial M$.

**Stokes’ Theorem:** Let $\mu$ be a $C^1$ $(n-1)$-dimensional form and let $M$ be an oriented smooth manifold with (possibly empty) oriented boundary $\partial M$ (with orientation induced by that of $M$). Then

$$\int_M d\mu = \int_{\partial M} \mu.$$ (333)

Outline proof: The proof of this theorem relies on the fundamental theorem of calculus. One first proves Stokes’ Theorem for the special case of the interior of an $n$-dimensional cube $[0,1]^n = I^n \subset \mathbb{R}^n$. (Strictly speaking this is not a manifold with boundary since the boundary itself has a boundary where it fails to be smooth, but this is on a set of measure zero on the boundary so can be ignored).

One can then extend this result to regions which are each the image of a cube under a diffeomorphism $f : U \to M$ where $I^n \subset U \subset \mathbb{R}^n$. Because both integration and exterior differentiation commute with the pullback Stokes’ Theorem is also true for $f(I^n)$.
One then writes $M$ as a sum of regions which are diffeomorphic to cubes. The contributions from the interior boundaries cancel and the contributions from the exterior boundaries add up to give the integral over $\partial M$.

It remains to prove the theorem for the case of $I^n$. We will illustrate the proof for the case of $I^2$. The general case is essentially the same but involves more complicated algebra to keep track of the boundary terms. Let

$$\mu = a(x, y)dx + b(x, y)dy$$

be a general 1-form on $I^2 \subset \mathbb{R}^2$. Then

$$d\mu = \frac{\partial a}{\partial y}dy \wedge dx + \frac{\partial b}{\partial x}dx \wedge dy$$

Then

$$\int_{I^2} d\mu = \int_{y=0}^{1} \int_{x=0}^{1} \left( \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) dx dy$$

Note: There is a formal mechanism for doing all these calculations based on the use of a chain complex. In our case the chain complex is based on $I^p$, but a more frequent choice is to base it on a $p$-simplex in which one has a simplicial complex. We will say more about this in the next Section.

5.4 The volume form on a Riemannian manifold

Let $(M, g)$ be a Riemannian manifold and let $x^i$ be a local coordinate system. Let $|g| = \det g$ be the determinant of $g_{ij}$. Then we define an $n$-form $\nu$ called the volume form defined by $g$ according to

$$\nu = |g|^\frac{1}{2} dx^1 \wedge \cdots \wedge dx^n$$

Now if we transform to a new coordinate system $\tilde{x}^i$,

$$\tilde{g}_{ij} = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial x^l}{\partial \tilde{x}^j} g_{kl}.$$
so that taking the determinant of both sides of the equation we obtain
\[
det \tilde{g} = det \left( \frac{\partial x^k}{\partial \tilde{x}^i} \right)^2 det g
\] (339)
and hence
\[
|\tilde{g}|^{\frac{1}{2}} = \det \left( \frac{\partial x^k}{\partial \tilde{x}^i} \right) |g|^{\frac{1}{2}}.
\] (340)
Therefore
\[
\tilde{\nu} = |\tilde{g}|^{\frac{1}{2}} d\tilde{x}^1 \wedge \cdots \wedge d\tilde{x}^n
\]
\[
= |g|^{\frac{1}{2}} \det \left( \frac{\partial x^k}{\partial \tilde{x}^i} \right) d\tilde{x}^1 \wedge \cdots \wedge d\tilde{x}^n
\]
\[
= \frac{1}{|g|^{\frac{1}{2}}} dx^1 \wedge \cdots \wedge dx^n
\]
\[
= \nu.
\] (341)
We have shown that the definition of \( \nu \) does not depend upon the coordinate system and defines a global \( n \)-form on \( M \).

Note that in a local coordinate system the volume form has components
\[
\nu_{i_1 \ldots i_n} = |g|^{\frac{1}{2}} \epsilon_{i_1 \ldots i_n}.
\] (342)

We may use the volume form to define the integral of a function on a Riemannian manifold by defining \( I = \int_M \phi \nu \). In the special case where we can cover \( M \) by a single coordinate system the integral of \( \phi \) is simply given by
\[
I = \int_{\mathbb{R}^n} \phi(x^1, \ldots, x^n) \sqrt{|g|} d^n x
\] (343)

5.5 The \( \ast \)-operator and the Hodge dual of a \( p \)-form

The contravariant form of the metric defines an inner product on 1-forms. Using the abstract index convention we may write this as
\[
g(\alpha, \beta) = g^{ab} \alpha_a \beta_b
\] (344)

We may extend this to an inner product on \( p \)-forms \( \alpha \) and \( \beta \) according to
\[
g(\alpha, \beta) = \frac{1}{p!} g^{a_1 b_1} g^{a_2 b_2} \cdots g^{a_p b_p} \alpha_{a_1 \ldots a_p} \beta_{b_1 \ldots b_p}
\] (345)

We now define a map (called the \textbf{star operator} or \textbf{Hodge dual}) from \( p \)-forms to \((n-p)\)-forms
\[
\ast : \Lambda^p(M) \rightarrow \Lambda^{n-p}(M)
\]
\[
\beta \mapsto \ast \beta
\] (346)
by requiring that
\[
\alpha \wedge \ast \beta = g(\alpha, \beta) \nu, \quad \forall \alpha \in \Lambda^p(M).
\] (347)
In components we therefore must have

\[ \beta_{a_{p+1} \ldots a_n} = \frac{1}{p!} \nu_{a_1 \ldots a_n} \beta_{b_1 \ldots b_p} g^{a_1 b_1} g^{a_2 b_2} \ldots g^{a_p b_p} \]

\[ = \frac{1}{p!} \nu_{a_1 \ldots a_n} \beta^{a_1 \ldots a_p} \]  

(348)

**Definition:** Let \( \xi \) be a vector field and \( \nu \) the volume form defined by the metric \( g \). Then \( \iota_\xi \nu \) is an \((n-1)\)-form and hence \( d(\iota_\xi \nu) \) is an \( n \)-form and must therefore be proportional to \( \nu \). We may therefore define the divergence \( \text{div} \xi \) by

\[ d(\iota_\xi \nu) = (\text{div} \xi) \nu. \]  

(349)

**Note:** if we use the metric to identify the vector field \( \xi \) with the 1-form \( \xi_i \) we can write \( \iota_\xi \nu \) as \( *\xi \) and the above equation as

\[ d *\xi = (\text{div} \xi) \nu. \]  

(350)

Taking the * of this equation we obtain

\[ *d *\xi = \text{div} \xi. \]  

(351)

We now show that this agrees with the standard definition of the divergence as \( \text{div} \xi = \nabla_i \xi^i \). Now

\[ \iota_\xi \nu = |g|^{\frac{1}{2}} \xi^i \epsilon_{j_1 \ldots j_n} dx^{j_1} \wedge \ldots \wedge dx^{j_n}. \]  

(352)

Hence

\[ d(\iota_\xi \nu) = \frac{\partial}{\partial x^{i_1}} \left( |g|^{\frac{1}{2}} \xi^i \right) \epsilon_{j_1 \ldots j_n} dx^{i_1} \wedge \ldots \wedge dx^{j_n} \]

\[ = \frac{\partial}{\partial x^j} \left( |g|^{\frac{1}{2}} \xi^i \right) \epsilon_{i_1 \ldots i_n} dx^{i_1} \wedge \ldots \wedge dx^{i_n}, \]  

(353)

so that

\[ \text{div} \xi = \frac{1}{|g|^{\frac{1}{2}}} \partial_i (|g|^{\frac{1}{2}} \xi^i). \]  

(354)

Since \( \Gamma^i_{ki} = \partial_k \ln |g|^{\frac{1}{2}} \) we can write this as

\[ \text{div} \xi = \nabla_i \xi^i, \]  

(355)

which is the usual definition.

**Divergence theorem:**

\[ \int_M (\text{div} \xi) \nu = \int_{\partial M} *\xi. \]  

(356)

**Proof:**

\[ \int_M (\text{div} \xi) \nu = \int_M d(\iota_\xi \nu) = \int_{\partial M} *\xi \]  

(357)

by Stokes’ Theorem.

### 5.6 Exercises
6 de Rham cohomology

6.1 Closed and exact forms

A differential form $\omega \in \Lambda^p(M)$ is said to be **closed** if $d\omega = 0$. The set of all closed $p$-forms is denoted $Z^p(M)$

$$Z^p(M) = \{ \omega \in \Lambda^p(M) : d\omega = 0 \}$$

(358)

If $\omega_1$ and $\omega_2$ are closed then so is $c_1\omega_1 + c_2\omega_2$ so that $Z^p(M)$ is a vector space over $\mathbb{R}$.

A differential form $\omega \in \Lambda^p(M)$ is said to be **exact** if $\omega = d\mu$ for some $\mu \in \Lambda^{p-1}(M)$. A linear combination of exact forms is exact so that $B^p(M)$ is also a vector space over $\mathbb{R}$. Furthermore since $d^2 = 0$ every exact from is closed so that $B^p(M)$ is a vector subspace of $Z^p(M)$. We will see later that locally the converse is also true so that every closed form may be locally written as an exact form (this is the content of Poincaré’s Lemma below) but that **globally** this is not true. The extent to which this fails to be the case encodes topological information about the manifold $M$ as shown by the following examples.

**Example 1:** Let $M = \mathbb{R}^2 \setminus \{(0,0)\}$ and consider the differential form

$$\omega = -\frac{y}{x^2 + y^2} \, dx + \frac{x}{x^2 + y^2} \, dy.$$  

(359)

A simple calculation shows that $d\omega = 0$, so that $\omega$ is a closed 1-form on $M$. However if $\omega$ were exact then we would have $\omega = d\phi$ for some function (0-form) $\phi$. If we then integrated $\omega$ over the unit circle $S$ we would have from Stokes’ theorem

$$\int_S \omega = \int_S d\mu = \int_{\partial S} \mu = 0,$$

since $\partial S = 0$.  

(360)

On the other hand a direct calculation shows that

$$\int_S \omega = 2\pi$$

(361)

so that $\omega$ cannot be exact. Note that locally $\omega = d\theta$ where $\theta = \tan^{-1} \frac{y}{x}$, but of course $\tan^{-1} \frac{y}{x}$ is not single-valued over the whole of $M$. This gives us an explicit example of a closed 1-form that is locally exact but not globally exact.

Going back to the general case suppose that $\omega \in \Lambda^p(M)$ is closed but not exact. Then if $\mu \in \Lambda^{p-1}(M)$, $\tilde{\omega} = \omega + d\mu$ is also an example of a differential form that is closed but not exact. Furthermore if $\omega_1$ and $\omega_2$ are (independent) differential forms that are closed but not exact then so is $c_1\omega_1 + c_2\omega_2$ ($c_1, c_2$ not both zero). From this we see that the key information is contained in the quotient space

$$H^p(M) = Z^p(M) / B^p(M).$$

(362)

Although both $Z^p(M)$ and $B^p(M)$ are infinite dimensional spaces, the quotient space turns out to be finite dimensional. In the example above with $M = \mathbb{R}^2 \setminus \{(0,0)\}$ we have seen that $H^1(M)$ is non-empty since it contains (the equivalence class of) the 1-form $\omega$ given by equation (359). Actually one can show that any closed but not exact 1-form $\tilde{\omega}$ on $M$ can be written as

$$\tilde{\omega} = c\omega + df, \quad \text{for } f \in \Lambda^0(M) \text{ and } c \in \mathbb{R},$$

(363)
so that in this case \( H^1(M) \) is one dimensional with basis \([\omega]\) the equivalence class of \(\omega\) in \(H^1(M)\).

If we consider \(M = \mathbb{R}^2 \setminus \{k \text{ distinct points}\}\) then each of the points gives rise to a closed but not exact 1-form \(\omega_k\) (given by a translated version of \(\omega\)) so that in this case \(H^1(M)\) is \(k\)-dimensional with basis \([\omega_1], \ldots, [\omega_k]\). Hence the dimension of \(H^1(M)\) tells us how many holes there are in the 2-dimensional manifold \(M\).

**Example 2a:** We now look at two 3-dimensional examples. If we look at \(M = \mathbb{R}^3 \setminus \{\text{the z-axis}\}\) then we still find that (359) is closed but not exact and that any such form can be written as \(c\omega + df\). Therefore, \(H^1(M)\) in this case is 1-dimensional and in general, if we look at \(\mathbb{R}^3 \setminus \{k \text{ distinct lines}\}\) then \(\dim H^1(M) = k\). Hence the dimension of \(H^1(M)\) tells us how many holes there are in the 2-dimensional manifold \(M\).

**Example 2b:** We now look at two 3-dimensional examples. If we look at \(M = \mathbb{R}^3 \setminus \{(0, 0, 0)\}\) and consider the 2-form
\[
\omega = \frac{zdx \wedge dy - ydx \wedge dz + xdy \wedge dz}{(x^2 + y^2 + z^2)^{3/2}},
\]
a direct calculation shows that \(d\omega = 0\), but that if we integrate \(\omega\) over the unit sphere we find \(\int_{S^2} \omega = 4\pi\) so by the Stokes’ Theorem argument we cannot have \(\omega\) exact since this would imply the integral over \(S^2\) vanished. In this example \(H^2(M)\) is generated by \([\omega]\) so that \(H^2(M)\) is 1-dimensional. More generally if we look at \(\mathbb{R}^3 \setminus \{k \text{ distinct points}\}\) then \(\dim H^2(M) = k\). We therefore see that in 3 dimensions \(\dim H^1\) and \(\dim H^2\) measure different sorts of holes in the space.

**Definition:**
\[
H^p(M) = Z^p(M)/B^p(M)
\]
is called the \(p\)-th de Rham cohomology group (note it is actually a vector space but may also be regarded as an abelian group under addition). Also,
\[
b^p = \dim H^p(M)
\]
is called the \(p\)-th Betti number on \(M\).

As we have seen the Betti numbers count the number of holes (of various dimensions) in the space \(\mathbb{R}^n\). We will prove later that if \(M\) and \(N\) are diffeomorphic then \(b^p(M) = b^p(N)\) so that the Betti numbers are topological invariants. We will make the relationship between cohomology and the topology of the manifold more precise later by briefly considering the homology of \(M\) and de Rham’s Theorem which relates the homology and the cohomology of \(M\).

**Example:** The Euler characteristic \(\chi = 2 - g\) (where \(g\) is the genus of \(M\)) may be given in terms of the Betti numbers
\[
\chi = \sum_{p=0}^{n} (-1)^p b^p.
\]

We now show that a closed form is locally exact. This is the content of the following result.

**Poincaré Lemma:** Let \(\omega \in Z^p(M)\) be a closed form. Then for every \(x \in M\) there exists an open neighbourhood \(V\) of \(x\) such that on \(V\) there exists \(\mu \in \Lambda^p(V)\) with \(\omega = d\mu\), i.e. every closed form is locally exact.

**Proof:** For any given \(x_0 \in M\) consider a chart \((\varphi, U)\) that maps \(x_0\) to the origin. Now let \(B\) be a ball with centre at the origin with sufficiently small
radius $r_0$ such that is in the image of $U$. Take $V = \varphi^{-1}(B)$. Since $\varphi^* \circ d = d \circ \varphi^*$ it is enough to prove Poincare’s Lemma for $B \in \mathbb{R}^n$.

In order to do this we construct a linear map $H : \Lambda^p(B) \to \Lambda^{p-1}(B)$ such that

$$H \circ d + d \circ H = \text{identity on } \Lambda^p(B).$$

(368)

Now if $d\omega = 0$ let $\mu = H\omega$. Then

$$\omega = (H \circ d + d \circ H)\omega = H(d\omega) + d(\omega H) = d\mu.$$

(369)

The required operator $H$ is defined by

$$H\omega(x) = \int_{t=0}^{1} t^{p-1}[t_x\omega](tx)dt, \quad x \in B \subset \mathbb{R}^n.$$

(370)

Note: This is only well-defined in $\mathbb{R}^n$ since we are thinking of $x \in \mathbb{R}^n$ as both the coordinate of a point (when we write $\omega(x)$) but also as the vector field $X$ with components $X^i(x) = x^i$ (when we write $t_x\omega$).

With $H\omega$ defined by (370) we have

$$[(H \circ d + d \circ H)\omega](x) = \int_{0}^{1} t^{p-1}[t_x \circ d + d \circ t_x \omega](tx)dt$$

$$= \int_{0}^{1} t^{p-1}[\mathcal{L}_x\omega](tx)dt.$$  

(371)

Now

$$(\mathcal{L}_x\omega)_{i_1...i_p} = x^j \omega_{i_1...i_p,j} + \sum_{k=1}^{p} x^j_{i_k} \omega_{i_1...j...i_p} = x^j \omega_{i_1...i_p,j} + \sum_{k=1}^{p} \delta^j_{i_k} \omega_{i_1...j...i_p} = x^j \omega_{i_1...i_p,j} + \omega_{i_1...i_p,j}.$$  

(372)

Hence

$$[(H \circ d + d \circ H)\omega](x) = \int_{0}^{1} t^{p-1}[\omega + x^j \omega_{j}](tx)dt$$

$$= \int_{0}^{1} \left( pt^{p-1}\omega(tx) + t^p \frac{\partial \omega}{\partial (tx^j)} \frac{d(tx^j)}{dt} \right) dt$$

$$= \int_{0}^{1} \frac{d}{dt}(t^p\omega(tx)) dt$$

$$= [t^p\omega(tx)]_0^1$$

$$= \omega(x),$$

(373)

as required.

**Proposition:** Let $f : M \to N$ be a smooth map between smooth manifolds $M$ and $N$.

If $\omega \in Z^p(N)$, then $f^*(\omega) \in Z^p(M)$.

If $\omega \in B^p(N)$, then $f^*(\omega) \in B^p(M)$.

58
Proof: This uses the fact that \( d \circ f^* = f^* \circ d \). If \( \omega \in Z^p(N) \) then \( d\omega = 0 \). But \( d(f^*\omega) = f^*(d\omega) = f^*(0) = 0 \), and hence \( f^*\omega \in Z^p(M) \).

On the other hand, if \( \omega \in B^p(N) \) then \( d\omega \neq 0 \). Hence \( f^*\omega = f^*d\mu = d(f^*\mu) \) and thus \( f^*\omega \in B^p(M) \).

Corollary: There exists a well defined map
\[
f^\# : H^p(N) \to H^p(M)
\]
defined by \( f^\#([\omega]) = [f^*\omega] \).

Theorem: The de Rham cohomology groups are invariant under diffeomorphisms \( f : M \to N \) between smooth manifolds. \( H^p(M) = f^\#(H^p(M)) \) and conversely \( H^p(N) = (f^{-1})^\#(H^p(M)) \).

Proof: \( f \circ f^{-1} = id \), hence \( (f \circ f^{-1})^* = id^* = id \) and thus
\[
(f^{-1})^* \circ f^* = id.
\]

Hence by the above corollary
\[
(f^{-1})^\# \circ f^\# = id.
\]
Similarly, \( f^{-1} \circ f = id \) proves that
\[
f^\# \circ (f^{-1})^\# = id.
\]

The above result shows that diffeomorphic manifolds have the same de Rham cohomology groups. In fact one can show a stronger result namely that manifolds of the same homotopy type have the same de Rham cohomology.

6.2 Integrals, the boundary operator \( \partial \) and homology

Let \( C \) be some smooth \((p+1)\)-dimensional region in \( M \) with \( p \)-dimensional boundary \( \partial C \) and \( \omega \) a \( p \)-form. Then Stokes’ Theorem tells us
\[
\int_{\partial C} \omega = \int_C d\omega.
\]

Writing the integral of a form as a pairing \( \langle \cdot, \cdot \rangle \) between the region and the form, we may write the above equation as
\[
\langle \partial C, \omega \rangle = \langle C, d\omega \rangle.
\]

If we now suppose that \( C \) is itself a boundary of some region \( D \), so that \( C = \partial D \), then \( \partial C = \partial^2 D \) and we may apply Stokes’ Theorem a second time to obtain
\[
\int_{\partial^2 D} \omega = \int_{\partial C} d\omega = \int_C d^2\omega = 0
\]
or using the \( \langle \cdot, \cdot \rangle \) notation we have
\[
\langle \partial^2 D, \omega \rangle = \langle \partial D, d\omega \rangle = \langle D, d^2\omega \rangle = 0
\]

Hence \( \partial^2 = 0 \) in the sense that \( \langle \partial^2 D, \omega \rangle = 0 \) for any choice of \( D \) and \( \omega \). There is therefore an analogy between the formula \( d^2 = 0 \) and \( \partial^2 = 0 \). This idea can be made precise using the concept of homology.
The concept of homology comes from applying the boundary operator $\partial$ to objects called $p$-chains. For simplicity in these notes we will consider chains built out of $p$-dimensional rectangles. These are simpler for the elementary calculations we will consider here but it is more usual to consider $p$-chains built out of simplexes. The concept of homology plays a key role in algebraic topology. In these notes we will simply present an elementary version suitable for our purposes.

Definition: An elementary $p$-chain $c$ on a manifold $M$ is a pair $c = (R, f)$ where $R$ is a $p$-dimensional rectangle

$$R = \{ (x^1, \ldots, x^p) \in \mathbb{R}^p : a^i \leq x^i \leq b^i \}$$

and a map

$$f : U \to M, \quad \text{where } U \text{ is an open subset of } \mathbb{R}^p \text{ containing } R$$

A $p$-chain $C$ is a formal linear combination of elementary $p$-chains

$$C = \sum_i \lambda_i c_i.$$  

The space of $p$-chains forms a vector space over $\mathbb{R}$.

Definition: Let $c$ be an elementary $p$-chain. We define its boundary $\partial c$ as the $p$-chain which is the sum of the 2$p$ boundary rectangles of $R$ (with a $+$ or $-$ sign according to orientation) and the map $f$ appropriately restricted to $\mathbb{R}^{p-1} \cap U$. Note the boundary of an elementary $p$-chain is the sum (with appropriate sign) of 2$p$ elementary $(p-1)$-chains. We now define the boundary of a $p$-chain to be the $(p-1)$-chain given by

$$\partial C = \sum_i \lambda_i \partial c_i.$$  

Definition: We may now define the integral of a $p$-form $\omega$ over a $p$-chain $C$ by

$$\int_C \omega = \sum_i \lambda_i \int_{C_i} \omega = \sum_i \lambda_i \int_{P_i \subset \mathbb{R}^p} f_i^* \omega.$$  

By the definition of the integral of a chain we may extend the pairing $\langle \cdot, \cdot \rangle$ to one between $p$-forms and $p$-chains. By the linearity of the integral we therefore still have (379) and (381), where now $C$ is a $(p+1)$-chain and $D$ a $(p+2)$-chain.

Definition: A finite chain $C$ such that $\partial C = 0$ is called a cycle. We denote the space of cycles by $Z_p(M)$.

$$Z_p(M) = \{ C : \partial C = 0 \}.$$  

A finite chain $C$ such that $C = \partial D$ is called a boundary. We denote the space of boundaries by

$$B_p(M) = \{ C : C = \partial D \}.$$  

Both $Z_p(M)$ and $B_p(M)$ are vector spaces. Since $\partial^2 = 0$ we the fact that $B_p$ is a vector subspace of $Z_p(M)$. We may define an equivalence relation on elements of $Z_p(M)$ by $C_1 \equiv C_2$ if $C_1 - C_2 = \partial D$ for some $(p+1)$-chain $D$. This enables us to define the $p$-th homology group of the chain complex by

$$H_p(M) = Z_p(M)/B_p(M).$$  

60
6.3 Homology and de Rham’s Theorem

In this section we establish a relationship between the homology groups and the de Rham cohomology groups on \( M \) using the pairing \( \langle \cdot, \cdot \rangle \) between chains and forms. We start with an example.

**Example:** Let \( U \) be a \((p+1)\)-dimensional region on \( M \) with \( p \)-dimensional boundary \( \partial U \) consisting of two parts \( S_1 \) and \( S_2 \), such that

\[
S_1 - S_2 = \partial U.
\]

(390)

Now let \( \omega \) be a closed \( p \)-form on \( M \). Then

\[
\int_{S_2} \omega - \int_{S_2} \omega = \int_{\partial U} \omega = \int_U d\omega = 0,
\]

(391)

so that

\[
\int_{S_1} \omega = \int_{S_2} \omega.
\]

(392)

We see that if \( S_1 \) and \( S_2 \) differ by a boundary term \( \partial U \), then the integrals over \( S_1 \) and \( S_2 \) are the same. Thus the integral of an exact form only depends upon the equivalence class \([S]\).

Similarly, if \( \omega_1 \) and \( \omega_2 \) are two \( p \)-forms that differ by \( d\mu \), or

\[
\omega_1 - \omega_2 = d\mu,
\]

(393)

then if we integrate the difference over a chain which is a cycle (ie \( \partial S = 0 \)), we obtain

\[
\int_S \omega_1 - \int_S \omega_2 = \int_S d\mu = \int_{\partial S} \mu = 0,
\]

(394)

so that the integral over a cycle only depends on the equivalence class \([\omega]\).

In this way we may elevate the pairing between \( p \)-chains and \( p \)-forms to one between equivalence classes.

**Proposition:** We may define a bilinear map

\[
\langle \cdot, \cdot \rangle : H_p(M) \times H^p(M) \to \mathbb{R}
\]

\[
([C], [\omega]) \mapsto \int_C \omega.
\]

(395)

**Proof:** Since the bilinearity is clear all we need to show that the above map is well defined. Suppose we choose different representatives \( \tilde{C} \in Z_p(M) \), with \( \tilde{C} = C + \partial D \) and \( \tilde{\omega} \in Z^p(M) \) with \( \tilde{\omega} = \omega + d\mu \). Then

\[
\int_{\tilde{C}} \tilde{\omega} = \int_{C} \tilde{\omega} + \int_{\partial D} \tilde{\omega}
\]

\[
= \int_C \tilde{\omega} + \int_{D} d\tilde{\omega}
\]

\[
= \int_C \tilde{\omega} \quad \text{(since} \ d\tilde{\omega} = 0\text{)}
\]

\[
= \int_C \omega + \int_{D} d\mu
\]

\[
= \int_C \omega + \int_{\partial C} \mu
\]

\[
= \int_C \omega.
\]

(396)
Hence
\[ \int_{\tilde{C}} \omega = \int_{C} \omega. \] (397)

**de Rham's theorem:** The bilinear map (395) is non-degenerate. i.e.
\[ \langle [C], \omega \rangle = 0 \quad \forall [C] \in H_{p}(M) \Rightarrow [\omega] = 0, \] (398)
\[ \langle [C], \omega \rangle = 0 \quad \forall [\omega] \in H^{p}(M) \Rightarrow [C] = 0. \] (399)

The proof of this result is quite long and beyond the scope of this course.

**Corollary:**
\[ \dim H^{p}(M) = \dim H_{p}(M). \] (400)

hence the Betti numbers are equal,
\[ b^{p}(M) = b_{p}(M). \] (401)

### 6.4 Exercises
7 Symplectic Geometry

7.1 Symplectic forms

**Definition:** A symplectic form on a manifold $M$ is a 2-form $\omega \in \Lambda^2(M)$ satisfying the conditions that it is:

(i) closed: $d\omega = 0$,

(ii) non-degenerate: at each point $x \in M$, $\omega(X,Y) = 0$, $\forall Y \in T_x M \Rightarrow X = 0$.

A symplectic manifold is a pair $(M,\omega)$ where $M$ is a manifold and $\omega$ a symplectic form.

**Remark:** Elementary linear algebra shows that a non-degenerate 2-form is possible only if $m = \dim M$ is even, so that $m = 2n$.

**Example 1:** Let $(x^1, \ldots, x^{2n})$ be the standard coordinates on $\mathbb{R}^{2n}$. Then

$$\omega = \sum_{i=1}^n dx^i \wedge dx^{n+i}$$

is a symplectic form on $\mathbb{R}^{2n}$. We will call this the standard symplectic form on $\mathbb{R}^{2n}$ and write it as $\omega$. One often writes $(x^1, \ldots, x^n) = (q^1, \ldots, q^n)$ and $(\tilde{x}^{n+1}, \ldots, \tilde{x}^{2n}) = (p_1, \ldots, p_n)$ (note that the index is deliberately down). One may then write

$$\tilde{\omega} = dq^i \wedge dp_i,$$

where the repeated index implies a summation over $i = 1 \ldots n$ (even though we are in $2n$ dimensions).

**Example 2:** Let $Q$ be a manifold and let $M = T^*Q$ be the cotangent bundle of $Q$. We will see that $T^*Q$ has a naturally defined 2-form which makes $(T^*Q,\omega)$ a symplectic manifold.

Let $q^i$ be local coordinates on $Q$. Then we may use the natural 1-forms $dq^i$ associated with the local coordinates $q^i$ as a basis for $T^*_qQ$ and write any 1-form $\alpha \in T^*_qM$ in the fibre above $q \in Q$ as $\alpha = p_i dq^i$, so that specifying $(q^i, p_i)$ gives local coordinates on $T^*Q$. We now define the symplectic form $\omega$ by

$$\omega = dq^i \wedge dp_i.$$  \hfill (404)

What does this 2-form look like in some other coordinate system? Let $\tilde{q}^i$ be some other system of local coordinates on $Q$. We now compute

$$\tilde{\omega} = d\tilde{q}^i \wedge d\tilde{p}_i.$$  \hfill (405)

Writing a general 1-form $\alpha$ in these coordinates and also in the original coordinates gives

$$\alpha = \tilde{p}_i d\tilde{q}^i = p_j dq^j = p_j \frac{\partial q^j}{\partial q^i} d\tilde{q}^i$$  \hfill (406)

Comparing the coefficients of $d\tilde{q}^i$ gives us

$$\tilde{p}_i = \frac{\partial q^j}{\partial \tilde{q}^i} p_j,$$  \hfill (407)

and hence

$$d\tilde{p}_i = \frac{\partial q^j}{\partial \tilde{q}^i} dp_j + p_j \frac{\partial^2 q^j}{\partial \tilde{q}^i \partial \tilde{q}^k} d\tilde{q}^k.$$  \hfill (408)
Hence
\[ \tilde{\omega} = dq^i \wedge dp_i \]
\[ = p_j \frac{\partial^2 q^i}{\partial q^j \partial q^k} dq^j \wedge dq^k + \frac{\partial q^i}{\partial q^j} dq^j \wedge \frac{\partial q^k}{\partial q^i} dp_k \]  
(409)
\[ = \frac{\partial q^i}{\partial q^j} dq^j \wedge dp_k \]
\[ = \delta^i_j dq^j \wedge dp_k \]
\[ = dq^j \wedge dp_j \]
\[ = \omega. \]  
(410)
(The first term in (410) vanishes because it is the contraction of a symmetric with an antisymmetric expression.) Hence the symplectic form does not depend on the coordinate system used.

There is in fact a coordinate-independent way of defining the symplectic form on \( T^*Q \). It is defined (with our sign convention) as minus the exterior derivative \( \omega = -d\theta \) of the canonical 1-form \( \theta \) on \( T^*Q \) (sometimes called the \textbf{Liouville 1-form}) defined below.

To define the canonical 1-form we first look at the projection map from \( T^*Q \) to \( Q \)
\[ \pi : T^*Q \rightarrow Q, \]  
(411)
with derivative
\[ \pi_* : T(T^*Q) \rightarrow TQ. \]  
(412)
A 1-form \( \theta \) on \( M = T^*Q \) gives a map \( \theta : TM \rightarrow \mathbb{R} \), so at each point \( x \in M \) we have a linear map \( \theta_x : T_xM \rightarrow \mathbb{R} \). Let \( x = (q, \alpha) \) be a point in \( M = T^*Q \) (where \( q \in Q \) and \( \alpha \in T^*_qQ \)). Now if \( V \in T_xM \), then \( (\pi)_x(V) \in T_qM \) and hence \( \alpha((\pi)_x(V)) \in \mathbb{R} \). Since everything is linear (and depends smoothly on \( x \)) we may define the required 1-form \( \theta \) on \( M = T^*Q \) by
\[ \theta_x(V) = \alpha_q((\pi)_x V). \]  
(413)
\textbf{Note:} If \( V = Q^i \frac{\partial}{\partial q^i} + P_i \frac{\partial}{\partial p_i} \) and \( \alpha = p_idq^i \) then \( (\pi)_x(V) = Q^i \frac{\partial}{\partial q^i} \) so that \( \alpha_q((\pi)_x V) = p_i Q^i \), and hence in local coordinates we have
\[ \theta = p_i dq^i, \]  
(414)
so that
\[ d\theta = dp_i \wedge dq^i = -\omega. \]  
(415)
as claimed.

Rather than consider further examples we first prove a theorem which shows that \textit{locally} all symplectic manifolds look like \((\mathbb{R}^{2n}, \omega)\) so that the interesting thing about symplectic manifolds is the global structure.

\textbf{Darboux’s Theorem:} For every point \( x_0 \) on a 2n-dimensional symplectic manifold \( (M, \omega) \) there exists an open neighbourhood \( U \) of \( x_0 \) and a chart
\[ \varphi : U \rightarrow \mathbb{R}^{2n}, \]  
(416)
such that
\[ \omega = \varphi^*(\tilde{\omega}), \]  
(417)
i.e. there exist local coordinates such that
\[ \omega = \sum_{i=1}^{n} dx^i \wedge dx^{n+i}. \] (418)

Equivalently there exist local coordinates so that
\[ \omega_{ij} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \] (419)

In order to prove the theorem we first state a useful lemma, the proof of which (exercise) follows directly from the definition of the Lie derivative.

**Lemma**: Let \( V_t(x) \) be a time-dependent ("velocity") vector field and \( \phi_t \) the corresponding flow i.e.
\[ \frac{d\phi_t}{dt} = V_t(\phi_t(x)). \] (420)

Then for any differential form \( \alpha \) we have
\[ \frac{d}{dt}(\phi_t^* \alpha) = \phi_t^*(L_{V_t} \alpha). \] (421)

**Proof of Darboux’s Theorem**: Let \( x_0 \) be some arbitrary point in \( M \) and let \( (U, \varphi) \) be a chart in some neighbourhood of \( x_0 \) with local coordinates \( x^i \).
Without loss of generality (after making some appropriate linear transformation of the coordinates) we may assume that the given form, which we will denote \( \omega_1 \) has the standard form at \( x_0 \). Let \( \omega_0 \) be the form which in the \( x^i \) coordinates has the standard symplectic form. Then our aim is to construct a diffeomorphism \( F: U \to F(U) \) keeping \( x_0 \) fixed and which is such that \( F^*_t(\omega_1) = \omega_0 \). Then in the chart given by \( \tilde{\varphi} = \varphi \circ F \circ \varphi^{-1} \) with local coordinates \( \tilde{x}^i \) the given form \( \omega_1 \) has the required canonical form.

To find such an \( F \) we first define \( \omega_t = (1-t)\omega_0 + t\omega_1 \) and construct a 1-parameter family of maps \( F_t \) such that
\[ F_t^* \omega_t = \omega_0. \] (422)

If the above is true we must have
\[ \frac{d}{dt}(F_t^* \omega_t) = 0, \] (423)
\[ \Rightarrow F_t^* \left( L_{V_t} \omega_t + \frac{d\omega_t}{dt} \right) = 0, \]
\[ \Rightarrow F_t^* \left( d_{V_t} \omega_t + \frac{d\omega_t}{dt} \right) = 0, \]
\[ \Rightarrow d_{V_t} \omega_t = \omega_0 - \omega_1. \] (424)

But \( d\omega_0 = 0 \) and \( d\omega_1 = 0 \) so that \( d(\omega_0 - \omega_1) = 0 \), so by Poincaré’s Lemma \( \omega_0 - \omega_1 = d\mu \) for some 1-form \( \mu \). We must therefore have
\[ d(\iota_{V_t} \omega_t - \mu) = 0. \] (425)

This will be satisfied if we choose \( V_t \) to satisfy the condition
\[ \iota_{V_t} \omega_t = \mu. \] (426)
Now since $\omega_t$ is non-degenerate (at least in a neighbourhood of $x_0$) then (426) defines a time dependent vector field. Let $F_t$ be the associated flow. Then we see that it will satisfy condition (423). Furthermore, since $\omega_t$ is constant (and equal to $\omega_0$) at the point $x_0$, we must also have $F_t(x_0) = x_0$ and

$$F_t^* \omega_t = \omega_0.$$  

(427)

$F_1$ is therefore the required transformation.

### 7.2 Hamiltonian Systems

In Hamiltonian mechanics the energy is written as a function of position $q^i$ and momentum $p_i$, or

$$H = H(q^i, p_i).$$  

(428)

The equations of motion are then given in these canonical coordinates by

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i},$$  

(429)

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}.$$  

(430)

**Example:** A particle of mass $m$ moving in 1-dimension under the influence of a potential $V(q)$ has kinetic energy $\frac{1}{2}mv^2 = \frac{1}{2m}p^2$ where $p = mv$ is the momentum. Its potential energy is $V(q)$ and hence the total energy is

$$H = \frac{1}{2m}p^2 + V(q).$$  

(431)

Hamilton’s equations are then

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{1}{m}p,$$  

(432)

$$\dot{p} = -\frac{\partial H}{\partial q} = -\frac{dV}{dq},$$  

(433)

or eliminating $p$,

$$m\ddot{q} = -\frac{dV}{dq}.$$  

(434)

**Note:** In more than one dimension, one needs to introduce the metric $\gamma_{ij}$ of Euclidean space to write the kinetic energy term (exercise).

We may use the language of symplectic geometry to write down Hamilton’s equations in a coordinate invariant way.

**Definition:** Let $(M, \omega)$ be a symplectic manifold and $H : M \to \mathbb{R}$ a function (called the “Hamiltonian”) on $M$, the $X_H$ is called the Hamiltonian vector field (associated to $H$) if

$$\iota_{X_H} \omega = dH.$$  

(435)

The triple $(M, \omega, X_H)$ is called a Hamiltonian system.

**Proposition:** When written in canonical coordinates (i.e. coordinates such that $\omega$ takes the standard form) a Hamiltonian vector field is given by

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}.$$  

(436)
for some Hamiltonian \( H \), and hence when written in these coordinates the flow equations are just Hamilton’s equations.

**Proof:** Let us take \( (q^i, p_i) \) to be canonical coordinates in which \( \omega \) takes the standard form \( \omega = dq^i \wedge dp_i \). Then if we write the components of a general Hamiltonian vector field to be

\[
X_H = Q^i \frac{\partial}{\partial q^i} + P_i \frac{\partial}{\partial p_i},
\]

\( i_{X_H} \omega \) is given by

\[
i_{X_H} \omega = Q^i dp_i - P_i dq^i.
\]

On the other hand

\[
dH = \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q^i} dq^i
\]

and hence in canonical coordinates a Hamiltonian vector field may be written as (436). If we now write the flow in components as \( f_t(x_0) = (q^i(t), p_i(t)) \) and substitute into the flow equation

\[
\frac{df_t}{dt} = X_H(f_t),
\]

we obtain (432, 433), i.e. the flow of a Hamiltonian vector field gives the solutions of Hamilton’s equations of motion.

**Theorem:** The Hamiltonian is preserved by the flow.

**Proof:**

\[
\frac{d}{dt} (H(f_t(x_0))) = DH_{f_t(x_0)}(f_t(x_0)) = dH(X_H) = \omega(X_H, X_H) = 0.
\]

**Liouville’s Theorem:** Let \( f_t \) be the flow of \( X_H \). Then \( f_t \) is symplectic, i.e.

\[
f_t^* \omega = \omega
\]

**Proof:** By the lemma (420) we have

\[
\frac{d}{dt} (f_t^* \omega) = f_t^* \mathcal{L}_{X_H} \omega = f_t^* (i_{X_H} d\omega + d(i_{X_H} \omega)) = f_t^* (0 + d(dH)) = 0.
\]

But \( f_0^* \omega = \omega \), so integration gives (442).

We may generalise Liouville’s Theorem to give a characterisation of local Hamiltonian vector fields: by this we mean a vector field \( X \) such that given any point \( x_0 \in M \) there is some open neighbourhood \( U \) and a real valued function \( H \) defined on \( U \) such that \( i_X \omega = dH \) on \( U \).

**Theorem:** The following conditions are equivalent:

1. \( X \) is a (local) Hamiltonian vector field for some \( H \).
that \( f^*_t \omega = \omega \) (where \( f_t \) is the flow generated by \( X \)).

(3) \( \mathcal{L}_X \omega = 0 \).

(4) \( \iota_X \omega \) is closed.

**Proof:** We have shown above that (1) \( \Rightarrow \) (2). Differentiating (2) and putting \( t = 0 \) gives (3). Since \( \mathcal{L}_X \omega = \iota_X d\omega + d(\iota_X \omega) \) we see that \( \mathcal{L}_X \omega = 0 \Rightarrow d(\iota_X \omega) = 0 \) and hence (3) \( \Rightarrow \) (4). Finally by using Poincaré’s Lemma we know that if \( \iota_X \omega \) is closed then locally we may write \( \iota_X \omega = dH \) so that \( X \) is a local Hamiltonian vector field. Hence (4) \( \Rightarrow \) (1).

**Proposition:** The set of locally Hamiltonian vector fields forms a Lie algebra with bracket given by the Lie bracket.

**Proof:**

\[
\mathcal{L}_{[X,Y]} \omega = \mathcal{L}_X \mathcal{L}_Y \omega - \mathcal{L}_Y \mathcal{L}_X \omega
\]

Hence if \( X \) and \( Y \) are two locally Hamiltonian vector fields then \( \mathcal{L}_X \omega = 0 \) and \( \mathcal{L}_Y \omega = 0 \) so by the above equation \( \mathcal{L}_{[X,Y]} = 0 \) and hence \([X,Y] \) is locally Hamiltonian.

The above result tells us that the Lie bracket of two locally Hamiltonian vector fields is locally Hamiltonian but does not tell us how to calculate the Hamiltonian. We end this section by introducing the Poisson bracket of two functions which will give us the required Hamiltonian.

**Definition:** Let \( \phi \) and \( \psi \) be real functions defined on a symplectic manifold \((M, \omega)\) then the Poisson bracket of \( \phi \) and \( \psi \) is given by

\[
\{ \phi, \psi \} = \omega(X_\phi, X_\psi),
\]

where \( X_\psi \) and \( X_\phi \) are the Hamiltonian vector fields generated by \( \phi \) and \( \psi \) respectively.

In canonical coordinates where the symplectic form takes the standard form then the Poisson bracket is given by (exercise)

\[
\{ \phi, \psi \} = \frac{\partial \phi}{\partial q^i} \frac{\partial \psi}{\partial p_i} - \frac{\partial \psi}{\partial q^i} \frac{\partial \phi}{\partial p_i}
\]

where we are using the summation convention over \( i = 1 \ldots n \).

**Proposition:** \( \{ \phi, H \} = 0 \) if and only if \( \phi \) is conserved by the flow.

**Proof:** In canonical coordinates we have

\[
\{ \phi, H \} = \frac{\partial \phi}{\partial p_i} \frac{\partial H}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial \phi}{\partial p_i}
\]

\[
= \frac{\partial \phi}{\partial q^i} \dot{q}^i + \frac{\partial \phi}{\partial p_i} \dot{p}_i
\]

\[
= \frac{d\phi}{dt}
\]

Hence \( \{ \phi, H \} = 0 \) if and only if \( \phi \) is conserved by the flow.

**Proposition:** The Poisson bracket satisfies the Jacobi identity

\[
\{\{ \phi, \psi \}, \chi \} + \{\psi, \chi \}, \phi \} + \{\chi, \phi \}, \psi \} = 0.
\]

We may use the above proposition to find new conserved quantities. If \( \phi \) and \( \psi \) are conserved then so is \( \{ \phi, \psi \} \). We may see this by taking \( h \) to be the Hamiltonian \( H \). Since \( \phi \) and \( \psi \) are conserved the second and third terms in
the Jacobi identity vanish and we are left with \( \{\{\phi, \psi\}, H\} = 0 \) which implies that \( \{\phi, \psi\} \) is conserved. The proposition below shows that the corresponding Hamiltonian vector field is given by the Lie bracket. Note unlike the result we had earlier for local Hamiltonian vector fields this is a global result.

**Proposition:**

\[ X_{\{\phi, \psi\}} = -[X_\phi, X_\psi]. \]  

(Proof by direct calculation.)

**Corollary:** The set of Hamiltonian vector fields generate a Lie algebra.

**Proof:** If \( X \) is a Hamiltonian vector field with Hamiltonian \( \phi \) and \( Y \) is a Hamiltonian vector field with Hamiltonian \( \psi \) then the above result shows that \( [X, Y] \) is a Hamiltonian vector field with Hamiltonian \( \{\phi, \psi\} \).

**Remark:** We can obtain from this result the fact that the set of Hamiltonian vector fields is a Lie subalgebra of the set of locally Hamiltonian vector fields with co-dimension \( \dim H^1(M) \), since for a local Hamiltonian vector field \( \iota_X \omega \) is closed, but for a global Hamiltonian vector field \( \iota_X \omega \) is exact.

We may use this result to show that if we have a Hamiltonian system with \( n \)-degrees of freedom and \( n \) independent conserved quantities \( H_1 \ldots H_n \) (including the Hamiltonian itself \( H = H_1 \)) then the system is integrable (see for example V I Arnold “Mathematical Methods of Mechanics” for details).

### 7.3 Exercises

1. Generalise the Hamiltonian theory of a point particle in a potential given in the text to \( n \) spatial dimension. You need to introduce the metric \( \gamma_{ij} \) of Euclidean space (or more generally, of configuration space).