Mathematical Methods of Physics III

Lecture Notes – Fall 2002

Claus Montonen – Esko Keski-Vakkuri

Contents

1 Introduction

The course Mathematical Methods of Physics III (MMP III) is third in the series of courses introducing mathematical concepts and tools which are often needed in physics. The first two courses MMP I-II focused on analysis, providing tools to analyze and solve the dynamics of physical systems. In MMP III the emphasis is on geometrical and topological concepts, needed for the understanding of the symmetry principles and topological structures of physics. In particular, we will learn group theory (the basic tool to understand symmetry in physics, especially useful in quantum mechanics, quantum field theory and beyond), topology (needed for many subtler effects in quantum mechanics and quantum field theory), and differential geometry (the language of general relativity and modern gauge field theories). There are also many more sophisticated areas of mathematics that are also often used in physics, notable omissions in this course are fibre bundles and complex geometry.

Course material will be available on the course homepage, to which you find a link from

`www.physics.helsinki.fi/∼tfo_www/lectures/courses.html`

Let me know of any typos and confusions that you find. The lecture notes often follow very closely (and often verbatim) the three recommended textbooks:

- H.F. Jones: Groups, Representations and Physics (IOP Publishing, 2nd edition, 1998)
- M. Nakahara: Geometry, Topology and Physics (IOP Publishing, 1990, a 2nd edition appeared in 2003, both editions will do)
- H. Georgi: Lie Algebras in Particle Physics (Addison-Wesley, 1982)

You don't necessarily have to rush to buy the books, they can be found in the reference section of the library in Physicum.

2 Group Theory

2.1 Group

Definition. A **group** G is a set of elements $\{a, b, \ldots\}$ with a law of composition (multiplication) which assigns to each ordered pair $a, b \in G$ another element $ab \in G$. (Note: $ab \in G$ (closure) is often necessary to check in order for the multiplication to be well defined). The multiplication must satisfy the following conditions:

G1 (associative law): For all $a, b, c \in G$, $a(bc) = (ab)c$.

G2 (unit element): There is an element $e \in G$ such that for all $a \in G$ $ae = ea = a$.

G3 (existence of inverse): For all $a \in G$ there is an element $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = e$.

If G satisfies **G1**, it is called a **semigroup**; if it also satisfies **G2**, it is called a **monoid**. The number of elements in the set G is called the **order** of the group, denoted by $|G|$. If $|G| < \infty$, G is a **finite** group. If G is a discrete set, G is a **discrete** group. If G is a continuous set, G is a **continuous** group.

Comments

i) In general $ab \neq ba$, i.e. the multiplication is not commutative. If $ab = ba$ for all $a, b \in G$, the group is called **Abelian**.

ii) The inverse element is unique: suppose that both b, b' are inverse elements of a. Then $b' = b'e = b'(ab) = (b'a)b = eb = b$.

Examples

1. Z with "+" (addition) as a multiplication is a discrete Abelian group.

2. R with "+" as a multiplication is a continuous Abelian group, $e = 0$. $R \setminus \{0\}$ with "$\cdot$" (product) is also a continuous Abelian group, $e = 1$. We had to remove 0 in order to ensure that all elements have an inverse.

3. $Z_2 = \{0, 1\}$ with addition modulo 2 is a finite Abelian group with order 2. $e = 0,\ 1^{-1} = 1$.

Let us also consider the set of mappings (functions) from a set X to a set Y, $Map(X,Y) = \{f : X \to Y | f(x) \in Y \ for\ all\ x \in X,\ f(x)\ is\ uniquely\ determined\}$. There are special cases of functions:

i) $f : X \to Y$ is called an **injection** (or **one-to-one**) if $f(x) \neq f(x')\ \forall x \neq x'$.

ii) $f : X \to Y$ is called a **surjection** (or **onto**) if $\forall y \in Y\ \exists x \in X\ s.t.\ f(x) = y$.

iii) if f is both an injection and a surjection, it is called a **bijection**.

Now take the composition of maps as a multiplication: $fg = f \circ g$, $(f \circ g)(x) = f(g(x))$. Then $(Map(X,X), \circ)$ (the set of functions $f : X \to X$ with $\circ$ as the multiplication) is a semigroup. We had to choose $Y = X$ to be able to use the composition, as g maps to Y but f is defined in X. Further, $(Map(X,X), \circ)$ is in fact a monoid with the identity map $id :\ id(x) = x$ as the unit element. However, it is *not* a group, unless we restrict to bijections. The set of bijections $f : X \to X$ is called the set of **permutations** of X, we denote $Perm(X) = \{f \in Map(X,X) | f\ is\ a\ bijection\}$. Every $f \in Perm(X)$ has an inverse map, so $Perm(X)$ is a group. However, in general $f(g(x)) \neq g(f(x))$, so $Perm(X)$ is not an Abelian group. An important special case is when X has a finite number N of elements. This is called the **symmetric group** or the **permutation group**, and denoted by S_N. The order of S_N is $|S_N| = N!$ (exercise).

Definitions

i) We denote $g^2 = gg,\ g^3 = ggg = g^2 g,\ \ldots,\ g^n = \overbrace{g \cdots g}^{n}$ for products of the element $g \in G$.

ii) The **order *n* of the element** $g \in G$ is the smallest number n such that $g^n = e$.

2.2 Smallest Finite Groups

Let us find all the groups of order n for $n = 1, \ldots, 4$. First we need a handy definition. A **homomorphism** in general is a mapping from one set X to another set Y preserving some structure. Further, if f is a bijection, it is called an **isomorphism**. We will see several examples of such structure-preserving mappings. The first one is the one that preserves the multiplication structure of groups.

Definition. A mapping $f : G \to H$ between groups G and H is called a **group homomorphism** if for all $g_1, g_2 \in G$, $f(g_1 g_2) = f(g_1)f(g_2)$. Further, if f is also a bijection, it is called a **group isomorphism**. If there exists a group isomorphism between groups G and H, we say that the groups are **isomorphic**, and denote $G \cong H$. Isomorphic groups have an identical structure, so they can be identified – there is only one abstract group of that structure.

Now let us move ahead to groups of order n.

Order $n = 1$. This is the trivial group $G = \{e\}$, $e^2 = e$.

Order $n = 2$. Now $G = \{e, a\}$, $a \neq e$. The multiplications are $e^2 = e$, $ea = ae = a$. For a^2, let's first try $a^2 = a$. But then $a = ae = a(aa^{-1}) = a^2 a^{-1} = aa^{-1} = e$, a contradiction. So the only possibility is $a^2 = e$. We can summarize this in the **multiplication table** or **Cayley table**:

	e	a
e	e	a
a	a	e

This group is called Z_2. You have already seen another realization of it: the set $\{0, 1\}$ with addition modulo 2 as the multiplication. Yet another realization of the group is $\{1, -1\}$ with product as the multiplication. This illustrates what was said before: for a given abstract group, there can be many ways to describe it. Consider one more realization: the permutation group $S_2 = Perm(\{1, 2\})$. Its elements are

$$
\begin{aligned}
e &= \begin{pmatrix} 1 & 2 \\ \downarrow & \downarrow \\ 1 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \\
a &= \begin{pmatrix} 1 & 2 \\ \downarrow & \downarrow \\ 2 & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} ,
\end{aligned}
$$

the arrows indicate how the numbers are permuted, we usually use the notation in the right hand side without the arrows. For products of permutations, the order in which they are performed is "right to left": we first perform the permutation on the far right, then continue with the next one to the left, and so one. This convention is inherited from that with composite mappings: (fg)(x)=f(g(x)). We can now easily show that S_2 is isomorphic with Z_2. Take e.g. $\{1, -1\}$ with the product as the realization of Z_2. Then we define the mapping $i : Z_2 \to S_2 :\ i(1) = e,\ i(-1) = a$. It is easy to see that i is a group homomorphism, and it is obviously a bijection. Hence it is an isomorphism, and $Z_2 \cong S_2$. There is only one abstract group of order 2.

Order $n = 3$**.** Consider now the set $G = \{e, a, b\}$. It turns out that there is again only one possible group of order 3. We can try to determine it by completing its multiplication table:

	e	a	b
e	e	a	b
a	a	?	?
b	b	?	?

First, guess $ab = b$. But then $a = a(bb^{-1}) = (ab)b^{-1} = bb^{-1} = e$, a contradiction. Try then $ab = a$. But now $b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}a = e$, again contradiction. So $ab = e$. Similarly, $ba = e$. Then, guess $a^2 = a$. Now $a = aaa^{-1} = aa^{-1} = e$, doesn't work. How about $a^2 = e$? Now $b = a^2b = a(ab) = ae = a$, doesn't work. So $a^2 = b$. Similarly, can show $b^2 = a$. Now we have worked out the complete multiplication table:

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Our group is actually called Z_3. We can simplify the notation and call $b = a^2$, so $Z_3 = \{e, a, a^2\}$. Z_3 and Z_2 are special cases of *cyclic groups* $Z_n = \{e, a, a^2, \ldots, a^{n-1}\}$. They have a single "generating element" a with order n: $a^n = e$. The multiplication rules are $a^p a^q = a^{p+q (mod\ n)}$, $(a^p)^{-1} = a^{n-p}$. Sometimes in the literature cyclic groups are denoted by C_n. One possible realization of them is by complex numbers, $Z_n = \{e^{\frac{2\pi ik}{n}} | k = 0, 1, \ldots\}$ with product as a multiplication. This also shows their geometric interpretation: Z_n is the symmetry group of rotations of a regular directed polygon with n sides (see H.F.Jones). You can easily convince yourself that $Z_n = \{0, 1, \ldots, n-1\}$ with addition modulo n is another realization.

Order $n = 4$**.** So far the groups have been uniquely determined, but we'll see that from order 4 onwards we'll have more possibilities. Let's start with a definition.

Definition. A **direct product** $G_1 \times G_2$ of two groups is the set of all pairs (g_1, g_2) where $g_1 \in G_1$ and $g_2 \in G_2$, with the multiplication $(g_1, g_2) \cdot (g_1', g_2') = (g_1 g_1', g_2 g_2')$. The unit element is (e_1, e_2) where e_i is the unit element of G_i $(i = 1, 2)$. It is easy to see that $G_1 \times G_2$ is a group, and its order is $|G_1 \times G_2| = |G_1||G_2|$.

Now we can immediately find at least one group of order 4: the direct product $Z_2 \times Z_2$. Denote $Z_2 = \{e, f\}$ with $f^2 = e$, and introduce a shorter notation for the pairs: $E = (e, e)$, $A = (e, f)$, $B = (f, e)$, $C = (f, f)$. We can easily find the multiplication table,

	E	A	B	C
E	E	A	B	C
A	A	E	C	B
B	B	C	E	A
C	C	B	A	E

The group $Z_2 \times Z_2$ is sometimes also called "Vierergruppe" and denoted by V_4.

There is another group of order 4, namely the cyclic group $Z_4 = \{e, a, a^2, a^3\}$. It is not isomorphic with $Z_2 \times Z_2$. (You can easily check that it has a different multiplication table.) It can be shown (exercise) that there are no other groups of order 4, just the above two.

Order $n \geq 5$. As can be expected, there are more possible non-isomorphic groups of higher finite order. We will not attempt to categorize them much further, but will mention some interesting facts and examples.

Definition. If H is a subset of the group G such that

i) $\forall\, h_1, h_2 \in H : \; h_1 h_2 \in H$

ii) $\forall\, h \in H : \; h^{-1} \in H$,

then H is called a **subgroup** of G. Note as a result of **i)** and **ii)**, every subgroup must include the unit element e of G.

Trivial examples of subgroups are $\{e\}$ and G itself. Other subgroups H are called **proper subgroups** of G. For those, $|H| \leq |G| - 1$.

Example. Take $G = Z_3$. Are there any proper subgroups? The only possibilities could be $H = \{e, a\}$ or $H = \{e, a^2\}$. Note that in order for H to be a group of order 2, it should be isomorphic with Z_2. But since $a^2 \neq e$ (because $a^3 = e$) and $(a^2)^2 = a^3 a = a \neq e$, neither is. So Z_3 has no proper subgroups.

2.2.1 More about the permutation groups S_n

It is worth spending some more time on the permutation groups, because on one hand they have a special status in the theory of finite groups (for a reason that I will explain later) and on the other hand they often appear in physics.

Let $X = \{1, 2, \ldots, n\}$. Denote a bijection of X by $p : X \to X,\ i \mapsto p(i) \equiv p_i$. We will now generalize our notation for the elements of S_n, you already saw it for S_2. We denote a $P \in S_n \equiv Perm(X)$ by

$$P = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} .$$

Recall that the multiplication rule for permutations was the composite operation, with the "right to left" rule. In general, the multiplication is not commutative:

$$PQ = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n \\ q_1 & q_2 & \cdots & q_n \end{pmatrix} \neq QP .$$

So, in general, S_n is not an abelian group. (Except S_2.) For example, in S_3,

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \tag{1}$$

but

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} , \tag{2}$$

which is not the same.

The identity element is

$$E = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

and the inverse of P is

$$P^{-1} = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix} .$$

An alternative and very useful way of writing permutations is the **cycle notation**. In this notation we follow the permutations of one label, say 1, until we get back to where we started (in this case back to 1), giving one **cycle**. Then we start again from a label which was not already included in the previously found cycle, and find another cycle, and so on until all the labels have been accounted for. The original permutation has then been decomposed into a certain number of *disjoint* cycles. This is best illustrated by an example. For example, the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

of S_4 decomposes into the disjoint cycles $1 \to 2 \to 4 \to 1$ and $3 \to 3$. Reordering the columns we can write it as

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 & | & 3 \\ 2 & 4 & 1 & | & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \end{pmatrix} .$$

In a cycle the bottom row is superfluous: all the information about the cycle (like $1 \to 2 \to 4 \to 1$) is already included in the order of the labels in the top row. So we can shorten the notation by simply omitting the bottom row. The above example is then written as

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = (124)(3) .$$

As a further abbreviation of the notation, we omit the 1-cycles (like (3) above), it being understood that any labels not appearing explicitly just transform into themselves. With the new shortened cycle notation, (1) reads

$$(23)(132) = (12) \tag{3}$$

and (2) reads as

$$(132)(23) = (13) . \tag{4}$$

In general, any permutation can always be written as the product of disjoint cycles. What's more, the cycles *commute* since they operate on different indices, hence the cycles can be written in any order in the product. In listing the individual permutations of S_n it is convenient to group them by cycle structure, i.e. by the number and length of cycles. For illustration, we list the first permutation groups S_n:

$n = 2$: $S_2 = \{E, (12)\}$.

$n = 3$: $S_3 = \{E, (12), (13), (23), (123), (132)\}$.

$n = 4$: $S_4 = \{E, (12), (13), (14), (23), (24), (34), (12)(34), (13)(24), (14)(23),$
$(123), (132), (124), (142), (134), (143), (234), (243),$
$(1234), (1243), (1324), (1342), (1423), (1432)\}$.

You can see that the notation makes it quite easy and systematic to write down all the elements in a concise fashion.

The simplest non-trivial permutations are the 2-cycles, which interchange two labels. In fact, *any permutation* can be built up from products of 2-cycles. First, an r-cycle can be written as the product of $r - 1$ overlapping 2-cycles:

$$(n_1 n_2 \ldots n_r) = (n_1 n_2)(n_2 n_3) \cdots (n_{r-1} n_r) .$$

Then, since any permutation is a product of cycles, it can be written as a product of 2-cycles. This allows us to classify permutations as "even" and "odd". First, a 2-cycle which involves just one interchange of labels is counted as odd. Then, a product of 2-cycles is even (odd), if there are an even (odd) number 2-cycles. Thus, an r-cycle is **even** (**odd**), if r is odd (even). (Since it is a product of $r-1$ 2-cycles.) Finally, a generic product of cycles is even if it contains an even number of odd cycles, otherwise it is odd. In particular, the identity E is even. This allows us to find an interesting subgroup of S_n, the **alternating group** A_n which consists of the *even* permutations of S_n. The order of A_n is $|A_n| = \frac{1}{2} \cdot |S_n|$. Hence A_n is a proper subgroup of S_n. Note that the odd permutations do not form a subgroup, since any subgroup must contain the identity E which is even.

To keep up a promise, we now mention the reason why permutation groups have a special status among finite groups. This is because of the following theorem (we state it without proof).

Theorem 2.1 (Cayley's Theorem) *Every finite group of order n is isomorphic to a subgroup of S_n.*

Thus, because of Cayley's theorem, in principle we know everything about finite groups if we know everything about permutation groups and their subgroups.

As for physics uses of finite groups, the classic example is their role in solid state physics, where they are used to classify general crystal structures (the so-called crystallographic point groups). They are also useful in classical mechanics, reducing the number of relevant degrees of freedom in systems of symmetry. We may later study an example, finding the vibrational normal modes of a water molecule. In addition to these canonical examples, they appear in different places and roles in all kinds of areas of modern physics.

2.3 Continuous Groups

Continuous groups have an uncountable infinity of elements. The **dimension** of a continuous group G, denoted $\dim G$, is the number of continuous real parameters (coordinates) which are needed to uniquely parameterize its elements. In the product $g'' = g'g$, the coordinates of g'' must be continuous functions of the coordinates of g and g'. (We will make this more precise later when we discuss topology. The above requirement means that the set of real parameters of the group must be a *manifold*, in this context called the *group manifold*.)

Examples.

1. The set of real numbers R with addition as the product is a continuous group; $\dim R = 1$. Simple generalization: $R^n = \{(r_1, \ldots, r_n) | r_i \in R,\ i = 1, \ldots, n\} = \overbrace{R \times \cdots \times R}^{n\ times}$, with product $(r_1, \ldots, r_n) \cdot (r'_1, \ldots, r'_n) = (r_1 + r'_1, \ldots, r_n + r'_n)$, $\dim R^n = n$.

2. The set of complex numbers C with addition as the product, $\dim C = 2$ (recall that we count the number of real parameters).

3. The set of $n \times n$ real matrices $M(n, R)$ with addition as the product, $\dim M(n, R) = n^2$. Note group isomorphism: $M(n, R) \cong R^{n^2}$.

4. $U(1) = \{z \in C | |z|^2 = 1\}$, with multiplication of complex numbers as the product. $\dim U(1) = 1$ since there's only one real parameter $\theta \in [0, 2\pi]$, $z = e^{i\theta}$. Note a difference with $U(1)$ and R: both have $\dim = 1$ but the group manifold of the former is the circle S^1 while the group manifold of the latter is the whole infinite x-axis. A generalization of $U(1)$ is $U(1)^n = \overbrace{U(1) \times \cdots \times U(1)}^{n\ times}$, $(e^{i\theta_1}, \ldots, e^{i\theta_n}) \cdot (e^{i\theta'_1}, \ldots, e^{i\theta'_n}) = (e^{i(\theta_1 + \theta'_1)}, \ldots, e^{i(\theta_n + \theta'_n)})$. The group manifold of $U(1)^n$ is an n-torus $\overbrace{S^1 \times \cdots \times S^1}^{n}$. Again, the n-torus is different from R^n: on the former it is possible to draw loops which cannot be smoothly contracted to a point, while this is not possible on R^n.

All of the above examples are actually examples of **Lie groups**. Their group manifolds must be *differentiable manifolds*, meaning that we can take smooth (partial) derivatives of the group elements with respect to the real parameters. We'll give a precise definition later – for now we'll just focus on listing further examples of them.

2.3.1 Examples of Lie groups

1. The group of general linear transformations $GL(n, R) = \{A \in M(n, R) |\ \det A \neq 0\}$, with matrix multiplication as the product; $\dim GL(n, R) = n^2$. While $GL(n, R)$, $M(n, R)$ have the same dimension, their group manifolds have a different structure. To parameterize the elements of $M(n, R)$, only one coordinate neighborhood is needed (R^{n^2} itself). The coordinates are the matrix entries a_{ij}:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} .$$

In $GL(n, R)$, the condition $\det A \neq 0$ removes a hyperplane (a set of measure zero) from R^{n^2}, dividing it into two disconnected coordinate regions. In each region, the entries a_{ij} are again suitable coordinates.

2. A generalization of the above is $GL(n, C) = \{n \times n$ complex matrices with non − zero determinant$\}$, with matrix multiplication as the product. This has $\dim GL(n, C) = 2n^2$. Note that $GL(n, R)$ is a (proper) subgroup of $GL(n, C)$. The following examples are subgroups of these two.

3. The group of special linear transformations $SL(n,R) = \{A \in GL(n,R) | \det A = 1\}$. It is a subgroup of $GL(n,R)$ since $\det(AB) = \det A \det B$. The dimension is $\dim SL(n,R) = n^2 - 1$.

4. The orthogonal group $O(n,R) = \{A \in GL(n,R) | A^T A = 1_n\}$, *i.e.* the group of orthogonal matrices. (1_n denotes the $n \times n$ unit matrix.) A^T is the **transpose** of the matrix A:

$$A^T = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix},$$

i.e. if $A = (a_{ij})$ then $A^T = (a_{ji})$, the rows and columns are interchanged. Let's prove that $O(n,R)$ is a subgroup of $GL(n,R)$:

a) $1_n^T = 1_n$ so the unit element $\in O(n,R)$

b) If A, B are orthogonal, then AB is also orthogonal: $(AB)^T(AB) = B^T A^T AB = B^T B = 1_n$.

c) Every $A \in O(n,R)$ has an inverse in $O(n,R)$: $(A^{-1})^T = (A^T)^{-1}$ so $(A^{-1})^T A^{-1} = (A^T)^{-1} A^{-1} = (AA^T)^{-1} = ((A^T)^T A^T)^{-1} = 1_n^{-1} = 1_n$.

Note that orthogonal matrices preserve the length of a vector. The length of a vector $\vec{v}$ is $\sqrt{v_1^2 + \cdots v_n^2} = \sqrt{\vec{v}^T \vec{v}}$. A vector $\vec{v}$ gets mapped to $A\vec{v}$, so its length gets mapped to $\sqrt{(A\vec{v})^T (A\vec{v})} = \sqrt{\vec{v}^T A^T A \vec{v}} = \sqrt{\vec{v}^T \vec{v}}$, the same. We can interpret the orthogonal group as the group of rotations in R^n.
What is the dimension of $O(n,R)$? $A \in GL(n,R)$ has n^2 independent parameters, but the orthogonality requirement $A^T A = 1_n$ imposes relations between the parameters. Let us count how many relations (equations) there are. The diagonal entries of $A^T A$ must be equal to one, this gives n equations; the entries above the diagonal must vanish, this gives further $n(n-1)/2$ equations. The same condition is then automatically satisfied by the "below the diagonal" entries, because the condition $A^T A = 1_n$ is symmetric: $(A^T A)^T = A^T A = (1_n)^T = 1_n$. Thus there are only $n^2 - n - n(n-1)/2 = n(n-1)/2$ free parameters. So $\dim O(n,R) = n(n-1)/2$.
Another fact of interest is that $\det A = \pm 1$ for every $A \in O(n,R)$. Proof: $\det(A^T A) = \det(A^T) \det A = \det A \det A = (\det A)^2 = \det 1_n = 1 \Rightarrow \det A = \pm 1$. Thus the group $O(n,R)$ is divided into two parts: the matrices with $\det A = +1$ and the matrices with $\det A = -1$. The former part actually forms a subgroup of $O(n,R)$, called $SO(n,R)$ (you can figure out why this is true, and not true for the part with $\det A$=-1). So we have one more example:

5. The group of special orthogonal transformations $SO(n,R) = \{A \in O(n,R) | \det A = 1\}$. $\dim SO(n,R) = \dim O(n,R) = n(n-1)/2$.

6. The group of unitary matrices (transformations) $U(n) = \{A \in GL(n, C) |\ A^\dagger A = 1_n\}$, where $A^\dagger = (A^*)^T = (A^T)^*$: $(A^\dagger)_{ij} = (A_{ji})^*$. Note that $(AB)^\dagger = B^\dagger A^\dagger$. These preserve the length of complex vectors $\vec{z}$. The length is defined as $\sqrt{z_1^* z_1 + \cdots z_n^* z_n} = \sqrt{\vec{z}^\dagger \vec{z}}$. Under A this gets mapped to $\sqrt{(A\vec{z})^\dagger A\vec{z}} = \sqrt{\vec{z}^\dagger A^\dagger A \vec{z}} = \sqrt{\vec{z}^\dagger \vec{z}}$. The unitary matrices are rotations in C^n. We leave it as an exercise to show that $U(n)$ is a subgroup of $GL(n, C)$, and $\dim U(n) = n^2$. Note that $U(1) = \{a \in C |\ a^* a = 1\}$, its group manifold is the unit circle S^1 on the complex plane.

7. The special unitary group $SU(n) = \{A \in U(n) |\ \det A = 1\}$. This is the complex analogue of $SO(n, R)$, and is a subgroup of $U(n)$. Exercise: $\dim SU(n) = n^2 - 1$. $U(n)$ and $SU(n)$ groups are important in modern physics. You will probably first become familiar with $U(1)$, the group of phase transformations in quantum mechanics, and with $SU(2)$, in the context of spin. Let's take a closer look at the latter. It's dimension is three. What does its group manifold look like? Let's first parameterize the $SU(2)$ matrices with complex numbers a, b, c, d:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} , \ A^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} .$$

Then

$$\begin{aligned} \det A &= ad - bc = 1 \\ A^\dagger A &= \begin{pmatrix} |a|^2 + |c|^2 & a^* b + c^* d \\ b^* a + d^* c & |b|^2 + |d|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \end{aligned}$$

Let's first assume $a \neq 0$. Then $b = -c^* d / a^*$. Substituting to the determinant condition gives $ad - bc = d(|a|^2 + |c|^2)/a^* = d/a^* = 1 \Rightarrow d = a^*$. Then $c = -b^*$. So

$$A = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} .$$

Assume then $a = 0$. Now $|c|^2 = 1$, $c^* d = 0 \Rightarrow d = 0$. Then $|c|^2 = |b|^2 = 1$. Write $b = e^{i\beta}, c = e^{i\gamma}$. Then $\det A = -bc = e^{i(\beta + \gamma + \pi)} = 1 \rightarrow \gamma = -\beta + (2n+1)\pi$. Then $c = e^{i\gamma} = e^{-i\beta} e^{i(2n+1)\pi} = -e^{-i\beta} = -b^*$. Thus

$$A = \begin{pmatrix} 0 & b \\ -b^* & 0 \end{pmatrix} .$$

Let us trade the two complex parameters with four real parameters x_1, x_2, x_3, x_4: $a = x_1 + ix_2$, $b = x_3 + ix_4$. Then A becomes

$$A = \begin{pmatrix} x_1 + ix_2 & x_3 + ix_4 \\ -x_3 + ix_4 & x_1 - ix_2 \end{pmatrix} .$$

The determinant condition $\det A = 1$ then turns into the constraint

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$$

for the four real parameters. This defines an unit 3-sphere. More generally, we define an **n-sphere** $S^n = \{(x_1, \ldots, x_{n+1}) \in R^{n+1} | \sum_{i=1}^{n+1} x_i^2 = 1\}$. The group manifold of $SU(2)$ is a three-sphere S^3. (And the group manifold of $U(1)$ was a 1-sphere S^1. As a matter of fact, these are the only Lie groups with n-sphere group manifolds.) The n-sphere is an example of so-called pseudospheres. We'll meet other examples in an exercise.

8. As an aside, note that $O(n, R), SO(n, R), U(n), SU(n)$ were associated with rotations in R^n or C^n, keeping invariant the lengths of real or complex vectors. One can generalize from real and complex numbers to quaternions and octonions, and look for generalizations of the rotation groups. This produces other examples of (compact) Lie groups, the $Sp(2n), G_2, F_4, E_6, E_7$ and E_8. The *symplectic* group $Sp(2n)$ plays an important role in classical mechanics, it is associated with canonical transformations in phase space. The other groups crop up in string theory.

2.4 Groups Acting on a Set

We already talked about the orthogonal groups as rotations, implying that the group acts on points in R^n. We should make this notion more precise. First, review the definition of a homomorphism from p. 4, then you are ready to understand the following

Definition. Let G be a group, and X a set. The **(left) action** of G on X is a homomorphism $L : G \to Perm(X)$, $G \ni g \mapsto L_g \in Perm(X)$. Thus, L satisfies $(L_{g_2} \circ L_{g_1})(x) = L_{g_2}(L_{g_1}(x)) = L_{g_2 g_1}(x)$, where $x \in X$. The last equality followed from the homomorphism property. We often simplify the notation and denote $gx \equiv L_g(x)$. Given such an action, we say that X is a **(left) G-space**. Respectively, the **right action** of G in X is a homomorphism $R : G \to Perm(X)$, $R_{g_2} \circ R_{g_1} = R_{g_1 g_2}$ (note order in the subscript!), $xg \equiv R_g(x)$. We then say that X is a **right G-space**.

Two (left) G-spaces X, X' can be identified, if there is a bijection $i : X \to X'$ such that $i(L_g(x)) = L'_g(i(x))$ where L, L' are (left) actions of G on X, X'. A mathematician would say this in the following way: the diagram

$$\begin{array}{ccc} X & \xrightarrow{i} & X' \\ L_g \downarrow & \searrow & \downarrow L'_g \\ X & \xrightarrow{i} & X' \end{array}$$

commutes, *i.e.* the map in the diagonal can be composed from the vertical and horizontal maps through either corner.

Definition. The **orbit** of a point $x \in X$ under the action of G is the set $O_x = \{L_g(x)| \ g \in G\}$. In other words, the orbit is the set of all points that can be reached from x by acting on it with elements of G. Let's put this in another way, by first introducing a useful concept.

Definition. An **equivalence relation** $\sim$ in a set X is a relation between points in a set which satisfies

i) $a \sim a$ (reflective) $\forall \ a \in X$

ii) $a \sim b \Rightarrow b \sim a$ (symmetric) $\forall \ a, b \in X$

iii) $a \sim b$ and $b \sim c \Rightarrow a \sim c$ (transitive) $\forall \ a, b, c \in X$

Given a set X and an equivalence relation $\sim$, we can partition X into mutually disjoint subsets called **equivalence classes**. An equivalence class $[a] = \{x \in X| \ x \sim a\}$, the set of all points which are equivalent to a under $\sim$. The element a (or any other element in its equivalence class) is called the **representative** of the class. Note that $[a]$ is not an empty set, since $a \sim a$. If $[a] \bigcap [b] \neq \emptyset$, there is an $x \in X$ s.t. $x \sim a$ and $x \sim b$. But then, by transitivity, $a \sim b$ and $[a] = [b]$. Thus, different equivalence classes must be mutually disjoint ($[a] \neq [b] \Rightarrow [a] \bigcap [b] = \emptyset$). The set of all equivalence classes is called the **quotient space** and denoted by $X/\sim$.

Example. Let n be a non-negative integer. Define an equivalence relation among integers $r, s \in Z$: $r \sim s$ if $r - s = 0 \ (mod \ n)$. (Prove that this indeed is an equivalence relation.) The quotient space is $Z/\sim = \{[0], [1], [2], \ldots, [n-1]\}$. Define the addition of equivalence classes: $[a] + [b] = [a + b(modn)]$. Then $Z/\sim$ with addition as a multiplication is a finite Abelian group, isomorphic to the cyclic group: $Z/\sim \cong Z_n$. (Exercise: prove the details.)

Back to orbits then. A point belonging to the orbit of another point defines an equivalence relation: $y \sim x$ if $y \in O_x$. The equivalence class is the orbit itself: $[x] = O_x$. Since the set X is partitioned into mutually disjoint equivalence classes, it is partitioned into mutually disjoint orbits under the action of G. We denote the quotient space by X/G. It may happen that there is only one such orbit, then $O_x = X \ \forall x \in X$. In this case we say that the action of G on X is **transitive**, and X is a **homogenous space**.

Examples.

1. $G = Z_2 = \{1, -1\}$, $X = R$. Left actions: $L_1(x) = x$, $L_{-1}(x) = -x$. Orbits: $O_0 = \{0\}$, $O_x = \{x, -x\}$ ($\forall \ x \neq 0$). The action is not transitive.

2. $G = SO(2,R)$, $X = R^2$. Parameterize

$$SO(2,R) \ni g = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} ,$$

and write

$$R^2 \ni x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} .$$

Left action:

$$L_g(x) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta\ x_1 - \sin\theta\ x_2 \\ \sin\theta\ x_1 + \cos\theta\ x_2 \end{pmatrix}$$

(rotate vector x counterclockwise about the origin by angle θ). Orbits are circles with radius r about the origin: $O_0 = \{0\}, O_{x\neq 0} = \{x \in R^2 |\ x_1^2 + x_2^2 = r^2\}$, $r = \sqrt{x_1^2 + x_2^2}$. The action is not transitive. $R^2/SO(2,R) = \{r \in R|\ r \geq 0\}$.

3. $G = GL(n,R)$, $X = R^n$. Left action: $L_A(x) = x'$ where $x'_i = \sum_{j=1}^n A_{ij}x_j$. There are two orbits: The orbit of the origin 0 is $O_0 = \{0\}$, all other points lie on the second orbit. So the action is not transitive.

2.4.1 Conjugacy classes and cosets

We can also let the group act on itself, *i.e.* take $X = G$. A simple way to define the left action of G on G is the *translation*, $L_g(g') = gg'$. Every group element belongs to the orbit of identity, since $L_g(e) = ge = g$. So $O_e = G$, the action is transitive. A more interesting way to define group action on itself is by *conjugation*.

Definition. Two elements g_1, g_2 of a group G are **conjugate** if there is an element $g \in G$ such that $g_1 = gg_2g^{-1}$. The element g is called the **conjugating element**.

We then take conjugation as the left action, $L_g(g') = gg'g^{-1}$. In general conjugation is not transitive. The orbits have a special name, they are called **conjugacy classes**.

It is also very interesting to consider the action of subgroups H of G on G. Define this time a *right* action of H on G by translation, $R_h(g) = gh$. If H is a proper subgroup, the action need not be transitive.

Definition. The orbits, or the equivalence classes

$$[g] = \{g' \in G|\ \exists h \in H\ s.t.\ g' = gh\} = \{gh|\ h \in H\}$$

are called **left cosets** of H, and usually they are denoted gH. The quotient space $G/H = \{gH|\ g \in G\}$ is the set of left cosets. (Similarly, we can define the left action $L_h(g) = hg$ and consider the right cosets Hg. Then the quotient space is denoted $H\backslash G$.)

Comments.

1. $ghH = gH$ for all $h \in H$.
2. If $g_1H = g_2H$, there is an $h \in H$ such that $g_2 = g_1h$ *i.e.* $g_1^{-1}g_2 \in H$.
3. There is a one-one correspondence between the elements of every coset and between the elements of H itself. The map $f_g : H \to gH$, $f_g(h) = gh$ is obviously a surjection; it is also an injection since $gh_1 = gh_2 \Rightarrow h_1 = h_2$. In particular, if H is finite, all the orders are the same: $|H| = |gH| = |g'H|$. This leads to the following theorem:

Theorem 2.2 (Lagrange's Theorem) *The order $|H|$ of any subgroup H of a finite group G must be a divisor of $|G|$: $|G| = n|H|$ where n is a positive integer.*

Proof. Under right action of H, G is partitioned into mutually disjoint orbits gH, each having the same order as H. Hence $|G| = n|H|$ for some n.

Corollary. If $p = |G|$ is a prime number, then $G \cong Z_p$.

Proof. Pick $g \in G$, $g \neq e$, denote the order of the element g by m. Then $H = \{e, g, \dots g^{m-1}\} \cong Z_m$ is a subgroup of G. But according to Lagrange's theorem $|G| = nm$. For this to be prime, $n = 1$ or $m = 1$. But $g \neq e$, so $m > 1$ so $n = 1$ and $|G| = |H|$. But then it must be $H = G$.

Definition. Let the group G act on a set X. The **little group** of $x \in X$ is the subgroup $G_x = \{g \in G |\ L_g(x) = x\}$ of G. It contains all elements of G which leave x invariant. It obviously contains the unit element e, you can easily show the other properties of a subgroup. The little group is also sometimes called the **isotropy group, stabilizer** or **stability group**.

Back to cosets. The set of cosets G/H is a G-space, if we define the left action $l_g : G/H \to G/H$, $l_g(g'H) = gg'H$. The action is transitive: if $g_1H \neq g_2H$, then $l_{g_1g_2^{-1}}(g_2H) = g_1H$. The inverse is also true:

Theorem 2.3 *Let group G act transitively on a set X. Then there exists a subgroup H such that X can be identified with G/H. In other words, there exists a bijection $i : G/H \to X$ such that the diagram*

$$\begin{array}{ccc} G/H & \xrightarrow{i} & X \\ l_g \downarrow & \searrow & \downarrow L_g \\ G/H & \xrightarrow{i} & X \end{array}$$

commutes.

Proof. Choose a point $x \in X$, denote its isotropy group G_x by H. Define a map $i : G/H \to X$, $i(gH) = L_g(x)$. It is well defined: if $gH = g'H$, then $g = g'h$ with some $h \in H$ and $L_g(x) = L_{g'h}(x) = L_{g'}(L_h(x)) = L_{g'}(x)$. It is an injection: $i(gH) = i(g'H) \Rightarrow L_g(x) = L_{g'}(x) \Rightarrow x = L_{g^{-1}}(L_{g'}(x)) = L_{g^{-1}g'}(x) \Rightarrow g^{-1}g' \in H \Rightarrow g' = gh \Rightarrow gH = g'H$. It is also a surjection: G acts transitively so for all $x' \in X$ there exists g s.t. $x' = L_g(x) = i(gH)$. The diagram commutes: $(L_g \circ i)(g'H) = L_g(L_{g'}(x)) = L_{gg'}(x) = i(gg'H) = (i \circ l_g)(g'H)$.

Corollary. A consequence of the proof is that the orbit of a point $x \in X$, O_x, can be identified with G/G_x since G acts transitively on any one of its orbits. Thus the orbits are determined by the subgroups of G, in other words the action of G on X is determined by the subgroup structure.

Example. $G = SO(3, R)$ acts on R^3, the orbits are the spheres $|x|^2 = x_1^2 + x_2^2 + x_3^2 = r^2$, *i.e.* S^2 when $r > 0$. Choose the point $x =$ north pole $= (0, 0, r)$ on every orbit $r > 0$. Its little group is

$$G_x = \left\{ \begin{pmatrix} A_{2\times 2} & 0 \\ 0 & 1 \end{pmatrix} \mid A_{2\times 2} \in SO(2, R) \right\} \cong SO(2, R) \ .$$

By Theorem 2.3 and its Corollary, $SO(3, R)/SO(2, R) = S^2$.

2.4.2 Normal subgroups and quotient groups

Since the quotient space G/H is constructed out of a group and its subgroup, it is natural to ask if it can also be a group. The first guess for a multiplication law would be

$$(g_1H)(g_2H) = g_1g_2H \ .$$

This definition would be well defined if the right hand side is independent of the labeling of the cosets. For example $g_1H = g_1hH$, so we then need $g_1g_2H = g_1hg_2H$ *i.e.* find $h' \in H$ s.t. $g_1g_2h' = g_1hg_2$. But this is not always true. We can circumvent the problem if H belongs to a particular class of subgroups, so called *normal* (also called *invariant, selfconjugate*) subgroups.

Definition. A **normal subgroup** H of G is one which satisfies $gHg^{-1} = \{ghg^{-1} | \, h \in H\} = H$ for all $g \in G$.

Another way to say this is that H is a normal subgroup, if for all $g \in G, h \in H$ there exists a $h' \in H$ such that $gh = h'g$.

Consider again the problem in defining a product for cosets. If H is a normal subgroup, then $g_1hg_2 = g_1(hg_2) = g_1(g_2h') = g_1g_2h'$ is possible. One can show that the above multiplication satisfies associativity, existence of identity (it is eH)

and existence of inverse $(gH)^{-1} = g^{-1}H$. Hence G/H is a group if H is a normal subgroup. When G/H is a group, it is called a **quotient group**.

Comments:

1. If H is a normal subgroup, its left and right cosets are the same: $gH = Hg$.

2. If G is Abelian, all of its subgroups are normal.

3. $|G/H| = |G|/|H|$ (follows from Lagrange's theorem).

Example. Consider $G = SU(2)$, $H = \{1_2, -1_2\} \cong Z_2$. $A1_2 = 1_2A$ for all $A \in SU(2)$, hence H is a normal subgroup. One can show that the quotient group $G/H = SU(2)/Z_2$ is isomorphic with $SO(3, R)$. This is an important result for quantum mechanics, we will analyze it more in a future problem set.

This is also an example of a *center*. A **center** of a group G is the set of all elements of $g' \in G$ which commute with every element $g \in G$. In other words, it is the set $\{g' \in G|\ g'g = gg'\ \forall g \in G\}$. You can show that a center is a normal subgroup, so the quotient of a group and its center is a group. The center of $SU(2)$ is $\{1_2, -1_2\}$.

We finish by showing another way of finding normal subgroups and quotient groups. Let the map $\mu : G_1 \to G_2$ be a group homomorphism. Its **image** is the set

$$Im\mu = \{g_2 \in G_2|\ \exists g_1 \in G_1\ s.t.\ g_2 = \mu(g_1)\}$$

and its **kernel** is the set

$$Ker\mu = \{g_1 \in G_1|\ \mu(g_1) = e_2\}\ .$$

In other words, the kernel is the set of all elements of G_1 which map to the unit element of G_2. You can show that $Im\mu$ is a subgroup of G_2, $Ker\mu$ a subgroup of G_1. Further, $Ker\mu$ is a normal subgroup: if $k \in Ker\mu$ then $\mu(gkg^{-1}) = \mu(g)e_2\mu(g^{-1}) = \mu(gg^{-1}) = \mu(e_1) = e_2$ *i.e.* $gkg^{-1} \in Ker\mu$. Hence $G_1/Ker\mu$ is a quotient group. In fact, it also isomorphic with $Im\mu$!

Theorem 2.4 *$G_1/Ker\mu \cong Im\mu$.*

Proof. Denote $K \equiv Ker\mu$. Define $i : G_1/K \to Im\mu$, $i(gK) = \mu(g)$. If $gK = g'K$ then there is a $k \in K$ s.t. $g = g'k$. Then $i(gK) = \mu(g) = \mu(g'k) = \mu(g')e_2 = i(g'K)$ so i is well defined. Injection: if $i(gK) = i(g'K)$ then $\mu(g) = \mu(g')$ so $e_2 = (\mu(g))^{-1}\mu(g') = \mu(g^{-1})\mu(g') = \mu(g^{-1}g')$ so $g^{-1}g' \in K$. Hence $\exists k \in K$ s.t. $g' = gk$ so $g'K = gK$. Surjection: i is a surjection by definition. Thus i is a bijection. Homomorphism: $i(gKg'K) = i(gg'K) = \mu(gg') = \mu(g)\mu(g') = i(gK)i(g'K)$. i is a homomorphism and a bijection, *i.e.* an isomorphism.

For example, our previous example $SU(2)/Z_2 \cong SO(3,R)$ can be shown this way, by constructing a surjective homomorphism $\mu : SU(2) \to SO(3,R)$ such that $Ker\mu = \{1_2, -1_2\}$.

3 Representation Theory of Groups

In the previous section we discussed the action of a group on a set. We also listed some examples of Lie groups, their elements being $n \times n$ matrices. For example, the elements of the orthogonal group $O(n,R)$ corresponded to rotations of vectors in R^n. Now we are going to continue along these lines and consider the action of a generic group on a (complex) vector space, so that we can represent the elements of the group by matrices. However, a vector space is more than just a set, so in defining the action of a group on it, we have to ensure that it respects the vector space structure.

3.1 Complex Vector Spaces and Representations

Definition. A complex **vector space** V is an Abelian group (we denote its multiplication by "+" and call it a sum), where an additional operation, **scalar multiplication** by a complex number $\mu \in C$ has been defined, such that the following conditions are satisfied:

i) $\mu(\vec{v}_1 + \vec{v}_2) = \mu\vec{v}_1 + \mu\vec{v}_2$

ii) $(\mu_1 + \mu_2)\vec{v} = \mu_1\vec{v} + \mu_2\vec{v}$

iii) $\mu_1(\mu_2\vec{v}) = (\mu_1\mu_2)\vec{v}$

iv) $1\ \vec{v} = \vec{v}$

v) $0\ \vec{v} = \vec{0}$ ($\vec{0}$ is the unit element of V)

We could have replaced complex numbers by real numbers, to define a real vector space, or in general replaced the set of scalars by something called a "field". Complex vector spaces are relevant for quantum mechanics. A comment on notations: we denote vectors with arrows: $\vec{v}$, but textbooks written in English often denote them in boldface: **v**. If it is clear from the context whether one means a vector or its component, one may also simply use the notation v for a vector.

Definition. Vectors $\vec{v}_1, \ldots, \vec{v}_n \in V$ are **linearly independent**, if $\sum_{i=1}^{n} \mu_i\vec{v}_i = \vec{0}$ only if the coefficients $\mu_1 = \mu_2 = \cdots = \mu_n = 0$. If there exist at most n linearly independent vectors, n is the **dimension** of V, we denote $\dim V = n$. If $\dim V = n$, a set $\{\vec{e}^1, \ldots, \vec{e}^n\}$ of linearly independent vectors is called a **basis** of the vector space. Given a basis, any vector $\vec{v}$ can be written in a form $\vec{v} = \sum_{i=1}^{n} v_i\vec{e}^i$, where the **components** v_i of the vector are found uniquely.

Definition. A map $L : V_1 \to V_2$ between two vector spaces V_1, V_2 is **linear**, if it satisfies

$$L(\mu_1 \vec{v}_1 + \mu_2 \vec{v}_2) = \mu_1 L(\vec{v}_1) + \mu_2 L(\vec{v}_2)$$

for all $\mu_1, \mu_2 \in C$ and $\vec{v}_1, \vec{v}_2 \in V$. A linear map is also called a **linear transformation**, or especially in physics context, a **(linear) operator**. If a linear map is also a bijection, it is called an **isomorphism**, then the vector spaces V_1 and V_2 are **isomorphic**, $V_1 \cong V_2$. It then follows that $\dim V_1 = \dim V_2$. Further, all n-dimensional vector spaces are isomorphic. An isomorphism from V to itself is called an **automorphism**. The set of automorphisms of V is denoted $\mathrm{Aut}(V)$. It is a group, with composition of mappings $L \circ L'$ as the law of multiplication. (Existence of inverse is guaranteed since automorphisms are bijections).

Definition. The **image** of a linear transformation is

$$\mathrm{im} L = f(V_1) = \{L(\vec{v}_1) |\ \vec{v}_1 \in V_1\} \subset V_2$$

and its **kernel** is the set of vectors of V_1 which map to the null vector $\vec{0}_2$ of V_2:

$$\ker L = \{\vec{v}_1 \in V_1 |\ L(\vec{v}_1) = \vec{0}_2\} \subset V_1 \ .$$

You can show that both the image and the kernel are vector spaces. I also quote a couple of theorems without proofs.

Theorem 3.1 $\dim V_1 = \dim \ker L + \dim \mathrm{im} L$.

Theorem 3.2 *A linear map $L : V \to V$ is an automorphism if and only if* $\ker L = \{\vec{0}\}$.

Note that a linear map is defined uniquely by its action on the basis vectors:

$$L(\vec{v}) = L(\sum_{i=1}^{n} v_i \vec{e}^i) = \sum_i v_i L(\vec{e}^i)$$

then we expand the vectors $L(\vec{e}^i)$ in the basis $\{\vec{e}^j\}$ and denote the components by L_{ji}:

$$L(\vec{e}^i) = \sum_j L_{ji} \vec{e}^j .$$

Now

$$L(\vec{v}) = \sum_i \sum_j v_i L_{ji} \vec{e}^j = \sum_j \left(\sum_i L_{ji} v_i \right) \vec{e}^j \ ,$$

so the image vector $L(\vec{v})$ has the components $L(\vec{v})_j = \sum_i L_{ji} v_i$. Let $\dim V_1 = \dim V_2 = n$. The above can be written in the familiar matrix language:

$$\begin{pmatrix} L(\vec{v})_1 \\ L(\vec{v})_2 \\ \vdots \\ L(\vec{v})_n \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{nn} \\ \vdots & & \ddots & \vdots \\ L_{n1} & \cdots & & L_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} .$$

We will often shorten the notation for linear maps and write $L\vec{v}$ instead of $L(\vec{v})$, and $L_1 L_2 \vec{v}$ instead of $L_1(L_2(\vec{v}))$. From the above it should also be clear that the group of automorphisms of V is isomorphic with the group of invertible $n \times n$ complex matrices:

$$\mathrm{Aut}(V) = \{L : V \to V |\ L \text{ is an automorphism}\} \cong GL(n, C) .$$

(The multiplication laws are composition of maps and matrix multiplication.)

Now we have the tools to give a definition of a representation of a group. The idea is that we define the action of a group G on a vector space V. If V were just a set, we would associate with every group element $g \in G$ a permutation $L_g \in Perm(V)$. However, we have to preserve the vector space structure of V. So we define the action just as before, but replace the group $Perm(V)$ of permutations of V by the group $\mathrm{Aut}(V)$ of automorphisms of V.

Definition. A (linear) representation of a group G in a vector space V is a homomorphism $D : G \to \mathrm{Aut}(V)$, $G \ni g \mapsto D(g) \in \mathrm{Aut}(V)$. The **dimension of the representation** is the dimension of the vector space $\dim V$.

Note:

1. D is a homomorphism: $D(g_1 g_2) = D(g_1) D(g_2)$.
2. $D(g^{-1}) = (D(g))^{-1}$.

We say that a representation D is **faithful** if $\mathrm{Ker} D = \{e\}$. Then $g_1 \neq g_2 \Rightarrow D(g_1) \neq D(g_2)$. Whatever the $\mathrm{Ker} D$ is, D is always a faithful representation of the quotient group $G/\mathrm{Ker} D$.

A mathematician would next like to classify all possible representations of a group. Then the first question is when two representations are the same (equivalent).

Definition. Let D_1, D_2 be representations of a group G in vector spaces V_1, V_2. An **intertwining operator** is a linear map $A : V_1 \to V_2$ such that the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{A} & V_2 \\ D_1(g) \downarrow & \searrow & \downarrow D_2(g) \\ V_1 & \xrightarrow{A} & V_2 \end{array}$$

commutes, *i.e.* $D_2(g)A = AD_1(g)$ for all $g \in G$. If A is an isomorphism (we then need $\dim V_1 = \dim V_2$), the representations D_1 and D_2 are **equivalent**. In other words, there then exists a **similarity transformation** $D_2(g) = AD_1(g)A^{-1}$ for all $g \in G$.

Example. Let $\dim V_1 = n$, $V_2 = C^n$. Thus any n-dimensional representation is equivalent with a representation of G by invertible complex matrices, the homomorphism $D_2 : G \to GL(n, C)$.

Definition. A **scalar product** in a vector space V is a map $V \times V \to C$, $(\vec{v}_1, \vec{v}_2) \mapsto \langle \vec{v}_1 | \vec{v}_2 \rangle \in C$ which satisfies the following properties:

i) $\langle \vec{v} | \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 \rangle = \mu_1 \langle \vec{v} | \vec{v}_1 \rangle + \mu_2 \langle \vec{v} | \vec{v}_2 \rangle$

ii) $\langle \vec{v} | \vec{w} \rangle = \langle \vec{w} | \vec{v} \rangle^*$

iii) $\langle \vec{v} | \vec{v} \rangle \geq 0$ and $\langle \vec{v} | \vec{v} \rangle = 0 \Leftrightarrow \vec{v} = \vec{0}$.

Given a scalar product, it is possible to normalize (*e.g.* by the Gram-Schmidt method) the basis vectors such that $\langle \vec{e}^i | \vec{e}^j \rangle = \delta^{ij}$. Such an **orthonormal** basis is usually the most convenient on to use. The **adjoint** $A^\dagger$ of an operator (linear map) $A : V \to V$ is the one which satisfies $\langle \vec{v} | A^\dagger \vec{w} \rangle = \langle A\vec{v} | \vec{w} \rangle$ for all $\vec{v}, \vec{w} \in V$.

Definition. An operator (linear map) $U : V \to V$ is **unitary** if $\langle \vec{v} | \vec{w} \rangle = \langle U\vec{v} | U\vec{w} \rangle$ for all $\vec{v}, \vec{w} \in V$. Equivalently, a unitary operator must satisfy $U^\dagger U = id_V = 1$. It follows that the corresponding $n \times n$ matrix must be unitary, *i.e.* an element of $U(n)$. Unitary operators form a subgroup $\mathrm{Unit}(V)$ of $\mathrm{Aut}(V) \cong GL(n, C)$.

Definition. An **unitary representation** of a group G is a homomorphism $D : G \to \mathrm{Unit}(V)$.

Definition. If U_1, U_2 are unitary representations of G in V_1, V_2, and there exists an intertwining isomorphic operator $A : V_1 \to V_2$ which preserves the scalar product, $\langle A\vec{v} | A\vec{w} \rangle_{V_2} = \langle \vec{v} | \vec{w} \rangle_{V_1}$ for all $\vec{v}, \vec{w} \in V_1$, the represenations are **unitarily equivalent**.

Example. Every n-dimensional unitary representation is unitarily equivalent with a representation by unitary matrices, a homomorphism $G \to U(n)$.

As always after defining a fundamental concept, we would like to classify all possibilities. The basic problem in group representation theory is to classify all unitary representations of a group, up to unitary equivalence.

3.2 Symmetry Transformations in Quantum Mechanics

We have been aiming at unitary representations in complex vector spaces because of their applications in Quantum Mechanics (QM). Recall that the set of all possible states of a quantum mechanical system is the Hilbert space $\mathcal{H}$, a complex vector space with a scalar product. State vectors are usually denoted by $|\psi\rangle$ as opposed to our previous notation $\vec{v}$, and the scalar product of two vectors $|\psi\rangle, |\chi\rangle$ is denoted $\langle\psi|\chi\rangle$. Note that usually the Hilbert space is an infinite dimensional vector space, whereas in our discussion of representation theory we've been focusing on finite dimensional vector spaces. Let's not be concerned about the possible subtleties which ensue, in fact in many cases finite dimensional representations will still be relevant, as you will see.

According to QM, the time evolution of a state is controlled by the Schrodinger equation,

$$\imath\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle$$

where H is the Hamilton operator, the time evolution operator of the system. Suppose that the system possesses a symmetry, with the symmetry operations forming a group G. In order to describe the symmetry, we need to specify how it acts on the state vectors of the system – we need to find its representation in the vector space of the states, the Hilbert space. The norm of a state vector, its scalar product with itself $\langle\psi|\psi\rangle$ is associated with a probability density and normalized to one, similarly the scalar product $\langle\psi|\chi\rangle$ of two states is associated with the probability (density) of measurements. Thus the representations of the symmetry group G must preserve the scalar product. In other words, the representations must be unitary. Moreover, in a closed system probability is preserved under the time evolution. Thus, unitarity of the representations must also be preserved under the time evolution.

We can summarize the above in a more formal way: if $g \mapsto U_g$ is a faithful unitary representation of a group G in the Hilbert space of a quantum mechanical system, such that for all $g \in G$

$$U_g H U_g^{-1} = H \tag{5}$$

where H is the Hamilton operator of the system, the group G is a **symmetry group** of the system.

The condition (5) arises as follows. Suppose a state vector $|\psi\rangle$ is a solution of the Schrodinger equation. In performing a symmetry operation on the system, the state vector is mapped to a new vector $U_g|\psi\rangle$. But if the system is symmetric, the new state $U_g|\psi\rangle$ must also be a solution of the Schrodinger equation: $i\hbar(d/dt)U_g|\psi\rangle = HU_g|\psi\rangle)$. But then it must be $i\hbar(d/dt)|\psi\rangle = i\hbar(d/dt)U_g^{-1}U_g|\psi\rangle = U_g^{-1}HU_g|\psi\rangle = H|\psi\rangle \Rightarrow U_g^{-1}HU_g = H$.

Consider in particular the energy eigenstates $|\phi_n\rangle$ at energy level E_n:

$$H|\phi_n\rangle = E_n|\phi_n\rangle\ .$$

An energy level may be degenerate, say with k linearly independent energy eigenstates $\{|\phi_{n1},\ldots,|\phi_{nk}\rangle\}$. They span a k-dimensional vector space $\mathcal{H}_n$, a subspace of the full Hilbert space. If the system is has a symmetry group,

$$HU_g|\phi_n\rangle = U_g H|\phi_n\rangle = E_n U_g|\phi_n\rangle$$

so all states $U_g|\phi_n\rangle$ are eigenstates at the same energy level E_n. Thus the representation U_g maps the eigenspace $\mathcal{H}_n$ to itself; in other words the representation U_g is a k-dimensional representation of G acting in $\mathcal{H}_n$. By an inverse argument, suppose that the system has a symmetry group G. Its representations then determine the possible degeneracies of the energy levels of the system.

3.3 Reducibility of Representations

It turns out that some representations are more fundamental than others. A generic representation can be decomposed into so-called irreducible representations. That is our next topic. Again, we start with some definitions.

Definition. A subset W of a vector space V is called a **subspace** if it includes all possible linear combinations of its elements: if $\vec{v},\vec{w}\in W$ then $\lambda\vec{v}+\mu\vec{w}\in W$ for all $\lambda,\mu\in C$.

Let D be a representation of a group G in vector space V. The representation space V is also called a **G-module**. (This terminology is used in Jones.) Let W be a subspace of V. We say that W is a **submodule** if it is closed under the action of the group G: $\vec{w}\in W \Rightarrow D(g)\vec{w}\in W$ for all $g\in G$. Then, the restriction of $D(g)$ in W is an automorphism $D(g)_W : W\to W$.

Definition. A representation $D: G\to \mathrm{Aut}(V)$ is **irreducible**, it the only submodules are $\{\vec{0}\}$ and V. Otherwise the representation is **reducible**.

Example. Choose a basis $\{\vec{e^i}\}$ in V, let $\dim V = n$. Suppose that all the matrices $D(g)_{ij} = \langle\vec{e^i}|D(g)ve^j\rangle$ turn out to have the form

$$D(g) = \begin{pmatrix} M(g) & S(g) \\ 0 & T(g) \end{pmatrix} \tag{6}$$

where $M(g)$ is a $n_1\times n_1$ matrix, $T(g)$ is a $n_2\times n_2$ matrix, $n_1+n_2=n$, and $S(g)$ is a $n_1\times n_2$ matrix. Then the representation is reducible, since

$$W = \left\{ \begin{pmatrix} \vec{v} \\ \vec{0} \end{pmatrix} \mid \vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_{n_1} \end{pmatrix} \right\} \tag{7}$$

is a submodule:

$$D(g)\begin{pmatrix}\vec{v}\\ \vec{0}\end{pmatrix}=\begin{pmatrix}M(g)\vec{v}+S(g)\vec{0}\\ T(g)\vec{0}\end{pmatrix}=\begin{pmatrix}M(g)\vec{v}\\ \vec{0}\end{pmatrix}\in W. \tag{8}$$

If in addition $S(g) = 0$ for all $g \in G$, the representation is obviously built up by combining two representations $M(g)$ and $T(g)$. It is then an example of a *completely reducible* representation. We'll give a formal definition shortly.

Definition. A **direct sum** $V_1 \oplus V_2$ of two vector spaces V_1 and V_2 consists of all pairs (v_1, v_2) with $v_1 \in V_1, v_2 \in V_2$, with the addition of vectors and scalar multiplication defined as

$$\begin{aligned}(v_1, v_2) + (v_1', v_2') &= (v_1 + v_1', v_2 + v_2')\\ \lambda(v_1, v_2) = (\lambda v_1, \lambda v_2)\end{aligned}$$

It is simple to show that $\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$. If a scalar product has been defined in V_1 and V_2, one can define a scalar product in $V_1 \oplus V_2$ by

$$\langle (v_1, v_2)|(v_1', v_2')\rangle = \langle v_1, v_1'\rangle + \langle v_2|v_2'\rangle \ .$$

Suppose D_1, D_2 are representations of G in V_1, V_2, one can then define a **direct sum** representation $D_1 \oplus D_2$ in $V_1 \oplus V_2$:

$$(D_1 \oplus D_2)(g)(v_1, v_2) = (D_1(g)v_1, D_2(g)v_2) \ .$$

In this case it is useful to adopt the notation

$$V_1 = \left\{\begin{pmatrix}\vec{v}_1\\ \vec{0}\end{pmatrix}\right\} \ ; \ V_2 = \left\{\begin{pmatrix}\vec{0}\\ \vec{v}_2\end{pmatrix}\right\}$$

so that

$$V_1 \oplus V_2 = \left\{\begin{pmatrix}\vec{v}_1\\ \vec{v}_2\end{pmatrix}\right\} = \{(\vec{v}_1, \vec{v}_2)\} \ .$$

Now the matrices of the direct sum representation are of the block diagonal form

$$(D_1 \oplus D_2)(g) = \begin{pmatrix}D_1(g) & 0\\ 0 & D_2(g)\end{pmatrix} \ .$$

Definition. A representation D in vector space V is **completely reducible** if for every submodule $W \subset V$ there exists a *complementary submodule* W' such that $V = W \oplus W'$ and $D \cong D_W \oplus D_{W'}$.

Comments.

1. According to the definition, we need to show that D is equivalent with the direct sum representation $D_W \oplus D_{W'}$. For the matrices of the representation, this means that there must be a similarity transformation which maps all the matrices $D(g)$ into a block diagonal form:

$$AD(g)A^{-1} = \begin{pmatrix} D_W(g) & 0 \\ 0 & D_{W'}(g) \end{pmatrix} .$$

2. Strictly speaking, according to the definition also an irreducible representation is completely reducible, as $W = V, W' = \{0\}$ or vice versa satisfy the requirements. We will exclude this case, and from now on by completely reducible representations we mean those which are not irreducible.

The goal in the **reduction** of a representation is to decompose it into irreducible pieces, such that

$$D \cong D_1 \oplus D_2 \oplus D_3 \oplus \cdots$$

(then $\dim D = \sum_i \dim D_i$). This is possible if D is completely reducible. So, given a representation, how do we know if it is completely reducible or not? Interesting representations from quantum mechanics point of view turn out to be completely reducible:

Theorem 3.3 *Unitary representations are completely reducible.*

Proof. Since we are talking about unitary representations, it is implied that the representation space V has a scalar product. Let W be a submodule. We define its *orthogonal complement* $W_\perp = \{\vec{v} \in V | \ \langle \vec{v}|\vec{w}\rangle = 0 \ \forall \vec{w} \in W\}$. I leave it as an excercise to show that $V \cong W \oplus W_\perp$. We then only need to show that $W_\perp$ is also a submodule (closed under the action of G). Let $\vec{v} \in W_\perp$, and denote the unitary representation by U. For all $\vec{w} \in W$ and $g \in G$ $\langle U(g)\vec{v}|\vec{w}\rangle = \langle U(g)\vec{v}|U(g)U^{-1}(g)\vec{w}\rangle = \langle \vec{v}|U^\dagger(g)U(g)U^{-1}(g)\vec{w}\rangle \stackrel{a}{=} \langle \vec{v}|U^{-1}(g)\vec{w}\rangle = \langle \vec{v}|U(g^{-1})\vec{w}\rangle \stackrel{b}{=} \langle \vec{v}|\vec{w}'\rangle \stackrel{c}{=} 0$, where the step a follows since U is unitary, step b since W is a G-module, and the step c is true since $\vec{v} \in W_\perp$. Thus $U(g)\vec{v} \in W_\perp$ so $W_\perp$ is closed under the action of G.

If G is a finite group, we can say more.

Theorem 3.4 *Let D be a finite dimensional representation of a finite group G, in vector space V. Then there exists a scalar product in V such that D is unitary.*

Proof. We can always define a scalar product in a finite dimensional vector space, *e.g.* by choosing a basis and defining $\langle\vec{v}|\vec{w}\rangle = \sum_{i=1}^{n} v_i^* w_i$ where v_i, w_i are the components of the vectors. Given a scalar product, we then define a "group averaged" scalar product $\langle\langle\vec{v}|\vec{w}\rangle\rangle = \frac{1}{|G|}\sum_{g'\in G}\langle D(g')\vec{v}|D(g')\vec{w}\rangle$. It is straightforward to show that $\langle\langle|\rangle\rangle$ satisfies the requirements of a scalar product. Further,

$$\begin{aligned}
\langle\langle D(g)\vec{v}|D(g)\vec{w}\rangle\rangle &= \frac{1}{|G|}\sum_{g'\in G}\langle D(g')D(g)\vec{v}|D(g')D(g)\vec{w}\rangle \\
&= \frac{1}{|G|}\sum_{g'\in G}\langle D(g'g)\vec{v}|D(g'g)\vec{w}\rangle \\
&= \frac{1}{|G|}\sum_{g''\in G}\langle D(g'')\vec{v}|D(g'')\vec{w}\rangle = \langle\langle\vec{v}|\vec{w}\rangle\rangle \ .
\end{aligned}$$

In other words, D is unitary with respect to the scalar product $\langle\langle|\rangle\rangle$.

Since we have previously shown that unitary representations are completely reducible, we have shown the following fact, called Maschke's theorem.

Theorem 3.5 (Maschke's Theorem) *Every finite dimensional representation of a finite group is completely reducible.*

3.4 Irreducible Representations

Now that we have shown that many representations of interest are completely reducible, and can be decomposed into a direct sum of irreducible representations, the next task is to classify the latter. We will first develop ways to identify inequivalent irreducible representations. Before doing so, we must discuss some general theorems.

Theorem 3.6 (Schur's Lemma) *Let D_1 and D_2 be two irreducible representations of a group G. Every intertwining operator between them is either a null map or an isomorphism; in the latter case the representations are equivalent, $D_1 \cong D_2$.*

Proof. Let A be an intertwining operator between the representations, *i.e.* the diagram

$$\begin{array}{ccc}
 & V_1 \overset{A}{\to} V_2 & \\
D_1(g)\downarrow & \searrow & \downarrow D_2(g) \\
 & V_1 \overset{A}{\to} V_2 &
\end{array}$$

commutes: $D_2(g)A = AD_1(g)$ for all $g \in G$. Let's first examine if A can be an injection. Note first that if $\mathrm{Ker}A \equiv \{\vec{v}\in V_1|\ A\vec{v} = \vec{0}_2\} = \{\vec{0}_1\}$, then A is an injection since if $A\vec{v} = A\vec{w}$ then $A(\vec{v}-\vec{w}) = 0 \Rightarrow \vec{v}-\vec{w} \in \mathrm{Ker}A = \{\vec{0}_1\} \Rightarrow \vec{v} = \vec{w}$. So what is $\mathrm{Ker}A$? Recall that $\mathrm{Ker}A$ is a subspace of V_1. Is it also a submodule, *i.e.* closed under the action of G? Let $\vec{v} \in \mathrm{Ker}A$. Then $AD_1(g)\vec{v} = D_2(g)A\vec{v} = \vec{0}_2$, hence $D_1(g)\vec{v} \in \mathrm{Ker}A$ *i.e.* $\mathrm{Ker}A$ is a submodule. But since D_1 is an irreducible representation, either $\mathrm{Ker}A = V_1$ or $\mathrm{Ker}A = \{\vec{0}_1\}$. In the former case all vectors of V_1 map to the null vector of V_2, so A is a null map $A = 0$. In the latter case, A is an injection. We then use a similar reasoning to examine if A is also a surjection. Let $\vec{v}_2 \in \mathrm{Im}A \equiv \{\vec{v} \in V_2 |\ \exists \vec{v}_1 \in V_1 \text{ s.t. } \vec{v} = A\vec{v}_1\}$. Then we can write $\vec{v}_2 = A\vec{v}_1$. Then $D_2(g)\vec{v}_2 = D_2(g)A\vec{v}_1 = A(D_1(g)\vec{v}_1)$ so also $D_2(g)\vec{v}_2 \in \mathrm{Im}A$. Thus, $\mathrm{Im}A$ is a submodule of V_2. But since D_2 is irreducible, either $\mathrm{Im}A = \{\vec{0}_2\}$ *i.e.* $A = 0$, or $\mathrm{Im}A = V_2$ *i.e.* A is a surjection. To summarize, either $A = 0$ or A is a bijection *i.e.* an isomorphism (since it is also a linear operator).

Corollary. If D is an irreducible representation of a group G in (complex) vector space V, then the only operator which commutes with all $D(g)$ is a multiple of the identity operator.

Proof. If $\forall g \in G\ AD(g) = D(g)A$, then for all $\mu \in C$ also $(A - \mu 1)D(g) = D(g)(A - \mu 1)$. According to Schur's lemma, either $(A - \mu 1)^{-1}$ exists for all $\mu \in C$or $(A - \mu 1) = 0$. However, it is always possible to find at least one $\mu \in C$ such that $(A - \mu 1)$ is not invertible. In the finite dimensional case this is follows from the fundamental theorem of algebra, which guarantees that the polynomial equation $\det(A - \mu 1) = 0$ has solutions for μ. (The infinite dimensional case is more delicate, but turns out to be true as well). So it must be $A = \mu 1$.

We will next discuss a sequence of theorems, starting from the rather abstract *fundamental orthogonality theorem* and then moving towards its more intuitive and user-friendly forms. Since we are interested in applications, I will cut some corners and skip the proof of the fundamental orthogonality theorem. It can be found in the literature (or in Montonen's handwritten notes) if you are interested in the details.

Theorem 3.7 (Fundamental Orthogonality Theorem) *Let U_1 and U_2 be two unitary irreducible representations of a group G in vector spaces V_1 and V_2. Then*

$$\sum_{g\in G} \langle \vec{w}_1 | U_1(g) \vec{v}_1 \rangle^*_{V_1} \langle \vec{w}_2 | U_2(g) \vec{v}_2 \rangle_{V_2} = \begin{cases} 0, \text{ if } U_1 \text{ and } U_2 \text{ are not equivalent} \\ \frac{|G|}{\dim V} \langle \vec{w}_1 | \vec{w}_2 \rangle^* \langle \vec{v}_1 | \vec{v}_2 \rangle, \text{ if } U_1 = U_2,\ V_1 = V_2 = V \end{cases}$$

for all $\vec{v}_1, \vec{w}_1 \in V_1$, $\vec{v}_2, \vec{w}_2 \in V_2$. In the latter case also $\dim V < \infty$.

Note that in the latter case $V_1 = V_2 = V$, so $\vec{v}_1, \vec{v}_2, \vec{w}_1, \vec{w}_2 \in V$ and the scalar products on the right hand side are those of V. While this is the generic form of the theorem, it is more insightful to consider a special case. In the latter case, pick an orthonormal basis $\{\vec{e}^i\}$ in V and choose $\vec{w}_1 = \vec{e}^i, \vec{v}_1 = \vec{e}^j$, $\vec{w}_2 = \vec{e}^k, \vec{v}_2 = \vec{e}^l$. Then, in the left hand side appear the matrices of the representation, $D^{(\alpha)}(g)_{kl} = \langle \vec{e}^k | U_\alpha(g) \vec{e}^l \rangle$

and the right hand side reduces to a product of Kronecker deltas. In other words, the FOGT takes the basis-dependent form

$$\sum_{g\in G} D_{ij}^{(\alpha)*}(g) D_{kl}^{(\beta)}(g) = \frac{|G|}{\dim D^{(\alpha)}}\ \delta_{\alpha\beta}\delta_{ik}\delta_{jl}\ . \tag{9}$$

The left hand side can be interpreted as a scalar product of two vectors, then the right hand side is an orthogonality relation for them. Namely, consider a given representation (labeled by α), and the ijth elements of its representation matrices. They form a $|G|$-component vector $(D_{ij}^{(\alpha)}(g_1), D_{ij}^{(\alpha)}(g_2), \ldots, D_{ij}^{(\alpha)}(g_{|G|}))$ where g_i are all the elements of the group G. So we have a collection of vectors, labeled by α, i, j. Then (9) is an orthogonality relation for the vectors, with respect to the scalar product $\langle \vec{v}|\vec{v}'\rangle = \sum_{i=1}^{|G|} v_i^* v_i'$. However, in a $|G|$ dimensional vector space there can be at most $|G|$ mutually orthogonal vectors. The index pair ij has $(\dim D^{(\alpha)})^2$ possible values, so the upper bound on the total number of the above vectors is

$$\sum_\alpha (\dim D^{(\alpha)})^2 \leq |G|\ ,$$

where the sum is taken over all possible unitary inequivalent representations (labeled by α). In fact (you can try to show it), the sum turns out to be equal to the order $|G|$. This theorem is due to Burnside:

Theorem 3.8 (Burnside's Theorem) $\sum_\alpha (\dim D^{(\alpha)})^2 = |G|$.

Burnside's theorem helps to rule out possibilities for irreducible representations. Consider *e.g.* $G = S_3$, $|S_3| = 6$. The possible dimensions of inequivalent irreducible representations are 2,1,1 or 1,1,1,1,1,1. It turns out that S_3 has only two inequivalent irreducible representations (show it). So the irreps have dimensions 2,1,1.

3.5 Characters

Characters are a convenient way to classify inequivalent irreducible representations.

To start with, let $\{\vec{e}^1, \ldots, \vec{e}^n\}$ be an orthonormal basis in a n-dimensional vector space V with respect to scalar product $\langle|\rangle$.

Definition. A **trace** of a linear operator A is

$$\mathrm{tr}\ A \equiv \sum_{i=1}^{n} \langle \vec{e}^i | A\vec{e}^i\rangle\ .$$

Note. Trace is well defined, since it is independent of a choice of basis. Let $\{\vec{e}'^1, \ldots, \vec{e}'^n\}$ be another basis. Then tr $A = \sum_i \langle \vec{e}^i | A\vec{e}^i \rangle = \sum_{ij} \langle \vec{e}^i | \vec{e}'^j \rangle \langle \vec{e}'^j | A\vec{e}^i \rangle = \sum_{ij} \langle A^\dagger \vec{e}'^j | \vec{e}^i \rangle \langle \vec{e}^i | \vec{e}'^j \rangle = \sum_{ij} \langle A^\dagger \vec{e}'^j | e'^j \rangle = \sum_j \langle \vec{e}'^j | A\vec{e}'^j \rangle$. Recall also that associated with the operator A is a $n \times n$ matrix with components $A_{ij} = \langle \vec{e}^i | A\vec{e}^j \rangle$. Thus tr A is equal to the trace of the matrix.

Now, let $D^{(\alpha)}(g)$ be an unitary representation of a finite group G in V.

Definition. The **character** of the representation $D^{(\alpha)}$ is the map

$$\chi^{(\alpha)} : G \to C,\ \chi^{(\alpha)}(g) = \text{tr } D^{(\alpha)}(g)\ .$$

Note. Equivalent representations have the same characters: tr $(AD^{(\alpha)}A^{-1}) = \text{tr } (A^{-1}AD^{(\alpha)}) = \text{tr } D^{(\alpha)}$, where we used cyclicity of the trace: tr $ABC = \text{tr } CAB = \text{tr } BCA$ etc.

Recall that conjugation $L_g(g_0) = gg_0g^{-1}$ is one way to define how G acts on itself, the orbits $\{gg_0g^{-1} |\ g \in G\}$ were called conjugacy classes. Since tr $D(gg_0g^{-1}) = \text{tr } (D(g)D(g_0)D^{-1}(g)) = \text{tr } D(g_0)$, group elements related by conjugation have the same character (again, use cyclicity of trace). So characters can be interpreted as mappings

$$\chi^{(\alpha)} : \{\text{conjugacy classes of } G\} \to C$$

Note also that the character of the unit element is the same as the dimension of the representation: $\chi^{(\alpha)}(e) = \text{tr } D^{(\alpha)}(e) = \text{tr } id_V = \dim V = \dim D^{(\alpha)}$.

Recall then the fundamental orthogonality theorem, in its basis-dependent form (9). Now we are going to set $i = j, k = l$ in (9) and sum over i and k. The left hand side becomes

$$\sum_{g \in G} \sum_i D^{(\alpha)*}_{ii}(g) \sum_k D^{(\beta)}_{kk}(g) = \sum_{g \in G} \chi^{(\alpha)*}(g) \chi^{(\beta)}(g)\ .$$

The right hand side becomes

$$\frac{|G|}{\dim D^{(\alpha)}}\ \delta_{\alpha\beta} \sum_{ik} \delta_{ik}\delta_{ik} = \frac{|G|}{\dim D^{(\alpha)}}\ \delta_{\alpha\beta} \sum_i \delta_{ii} = |G|\ \delta_{\alpha\beta}\ .$$

We have derived an **orthogonality theorem for characters**:

$$\sum_{g \in G} \chi^{(\alpha)*}(g) \chi^{(\beta)}(g) = |G|\ \delta_{\alpha\beta}\ . \tag{10}$$

It can be used to analyze the reduction of a representation. In the reduction of a representation D, it may happen that an irreducible representation $D^{(\alpha)}$ appears multiple times in the the direct sum:

$$D = D^{(1)} \oplus D^{(1)} \oplus D^{(1)} \oplus D^{(2)} \oplus D^{(3)} \oplus \cdots$$

Then we shorten the notation and multiply each irreducible representation by an integer n_α to account for how many times $D^{(\alpha)}$ appears:

$$D = 3D^{(1)} \oplus D^{(2)} \oplus D^{(3)} \oplus \cdots = \bigoplus_\alpha n_\alpha D^{(\alpha)} \ .$$

n_α is called the **multiplicity** of the representation $D^{(\alpha)}$ in the decomposition. Since tr is a linear operation, obviously the characters of the representation satisfy

$$\chi = \sum_\alpha n_\alpha \chi^{(\alpha)}$$

with the same coefficients n_α. If we know the character χ of the reducible representation D, and all the characters $\chi^{(\alpha)}$ of the irreducible representations, we can calculate the multiplicities of each irreducible representation in the decomposition by using the orthogonality theorem of characters:

$$n_\alpha = \frac{1}{|G|} \sum_g \chi^{(\alpha)*}(g)\chi(g) \ .$$

Then, once we know all the multiplities, we know what is the decomposition of the representation D. In practise, characters of finite groups can be looked up from *character tables*. You can find them *e.g.* in *Atoms and Molecules*, by M. Weissbluth, pages 115-125. For more explanation of construction of character tables, see Jones, section 4.4. You will work out some character tables in a problem set.

Again, the orthogonality of characters can be interpreted as an orthogonality relation for vectors, with useful consequences. Let $C_1, C_2, \ldots, C_k$ be the conjugacy classes of G, denote the number of elements of C_i by $|C_i|$. Then (10) implies

$$\sum_{\{C_i\}} |C_i| \chi^{(\alpha)*}(C_i)\chi^{(\beta)}(C_i) = |G| \ \delta_{\alpha\beta} \ . \tag{11}$$

Consider then the vectors $\vec{v}_\alpha = (\sqrt{|C_1|}\chi^{(\alpha)}(C_1), \ldots, \sqrt{|C_k|}\chi^{(\alpha)}(C_k))$. The number of such vectors is the same as the number of irreducible representations. On the other hand, (11) tells that the vectors are mutually orthogonal, so the can be no more of them than the dimension of the vector space k, the number of conjugacy classes. Again, it can be show that the numbers are actually the same:

Theorem 3.9 *The number of unitary irreducible representations of a finite group is the same as the number of its conjugacy classes.*

If the group is Abelian, the conjugacy class of each element contains only the element itself: $gg_0g^{-1} = g_0gg^{-1} = g_0$. So the number of conjugacy classes is the same as the order of the group $|G|$, this is then also the number of unitary irreducible representations. On the other hand, according to Burnside's theorem,

$$\sum_{\alpha=1}^{|G|} (\dim D^{(\alpha)})^2 = |G| \ .$$

Since there are $|G|$ terms on the left hand side, it must be $\dim D^{(\alpha)} = 1$ for all α. Hence:

Theorem 3.10 *All unitary irreducible representations of an Abelian group are one dimensional.*

This fact can be shown to be true even for continuous Abelian groups. (Hence no word "finite" in the above.)

4 Differentiable Manifolds

4.1 Topological Spaces

The **topology** of a space X is defined via its open sets.
Let X= set, $\tau = \{X_\alpha\}_{\alpha \in I}$ a (finite or infinite) collection of subsets of X. (X, τ) is a **topological space**, if

T1 $\emptyset \in \tau,\ X \in \tau$

T2 <u>all</u> possible unions of X_α's belong to τ $\left(\bigcup_{\alpha \in I'} X_\alpha \in \tau,\ I' \subseteq I\right)$

T3 all intersections of a <u>finite</u> number of X_α's belong to τ. $(\bigcap_{i=1}^{n} X_{\alpha_i} \in \tau)$

The X_α are called the **open sets** of X in topology τ, and τ is said to give a **topology** to X.
So: topology $\hat{=}$ specify which subsets of X are open.
The same set X has several possible definitions of topologies (see examples).

Examples

(i) $\tau = \{\emptyset, X\}$ "trivial topology"

(ii) $\tau = \{\text{all subsets of } X\}$ "discrete topology"

(iii) Let $X = \mathbb{R}$, $\tau = \{\text{open intervals }]a, b[\text{ and their unions}\}$ "usual topology"

(iv) $X = \mathbb{R}^n$, $\tau = \{\ ]a_1, b_1[\times \ldots \times]a_n, b_n[\text{ and unions of these.}\}$

Definition: A **metric** on X is a function $d : X \times X \to \mathbb{R}$ such that

M1 $d(x,y) = d(y,x)$

M2 $d(x,y) \geq 0$, and $d(x,y) = 0$ if and only if $x = y$.

M3 $d(x,y) + d(y,z) \geq d(x,z)$ "triangle inequality"

Example:

$$X = \mathbb{R}^n, \quad d_p(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}, \ p > 0$$

If $p = 2$ we call it the Euclidean metric.

If X has a metric, then the **metric topology** is defined by choosing all the "open disks"

$$U_\epsilon(x) = \{ \ y \in X | \ d(x,y) < \epsilon \ \}$$

and all their unions as open sets.
The metric topology of $\mathbb{R}^n$ with metric d_p is equivalent with the usual topology (for all $p > 0$!)

Let (X, τ) be a topological space, $A \subset X$ a subset. The topology τ induces the **relative topology** τ' in A,

$$\tau' = \{ \ U_i \cap A \mid U_i \in \tau \ \}$$

This is how we obtain a topology for all subsets of $\mathbb{R}^n$ (like S^n).

4.1.1 Continuous Maps

Let (X, τ) and (Y, σ) be topological spaces. A map $f : X \to Y$ is **continuous** if the inverse image of every open set $V \in \sigma$, $f^{-1}(V) = \{ \ x \in X \mid f(x) \in V\}$, is an open set in X: $f^{-1}(V) \in \tau$.

A function $f : X \to Y$ is a **homeomorphism** if f is continuous, and has an inverse $f^{-1} : Y \to X$ which is also continuous.

If there exists a homeomorphism $f : X \to Y$, then we say that X is **homeomorphic** to Y and vice versa. Denote $X \approx Y$.
This ($\approx$) is an equivalence relation.

Intuitively : X and Y are homeomorphic if we can continuously deform X to Y (without cutting or pasting).
Example: coffee cup $\approx$ doughnut.

[The fundamental question of topology : classify all homeomorphic spaces.]

One method of classification: **topological invariants** i.e. quantities which are invariant under homeomorphisms.
If a topological invariant for $X_1 \neq$ for X_2 then $X_1 \not\approx X_2$.

The **neighbourhood** N of a point $x \in X$ is a subset $N \subset X$ such that there exists an open set $U \in \tau,\ x \in U$ and $U \subset N$.
(N does not have to be an open set).

(X, τ) is a **Hausdorff** space if for an arbitrary pair $x, x' \in X,\ x \neq x'$, there always exists neighbourhoods $N \ni x,\ N' \ni x'$ such that $N \cap N' = \emptyset$.
We'll assume from now on that all topological spaces (that we'll consider) are Hausdorff.

Example: $\mathbb{R}^n$ with the usual topology is Hausdorff.
All spaces X with metric topology are Hausdorff.

A subset $A \subset X$ is **closed** if its complement $X - A = \{x \in X \mid x \notin A\}$ is open.
N.B. X and $\emptyset$ are both open and closed.

A collection $\{A_i\}$ of subsets $A_i \subset X$ is called a **covering** of X if $\bigcup_i A_i = X$.
If all A_i are open sets in the topology τ of X, $\{A_i\}$ is an **open covering**.

A topological space (X, τ) is **compact** if, for every open covering $\{U_i \mid i \in I\}$ there exists a finite subset $J \subset I$ such that $\{U_i \mid i \in J\}$ is also a covering of X, i.e. every open covering has a finite subcovering.

X is **connected** if it cannot be written as $X = X_1 \bigcup X_2$, with X_1, X_2 both open, nonempty and disjoint, i.e. $X_1 \bigcap X_2 = \emptyset$.

A loop in topological space X is a continuous map $f : [0, 1] \to X$ such that $f(0) = f(1)$. If any loop in X can be continuously shrunk to a point, X is called **simply connected**.

Examples: $\mathbb{R}^2$ is simply connected.

The torus T^2 is not simply connected.

Examples of topological invariants = quantities or properties invariant under homeomorphisms:

1. Connectedness
2. Simply connectedness
3. Compactness
4. Hausdorff
5. Euler characteristic (see below)

Let $X \subset \mathbb{R}^3$, $X \approx$ polyhedron K. (monitahokas)
Euler characteristic:

$$\chi(X) = \chi(K) = \text{ (\# vertices in } K) - (\text{\# edges in } K) + (\text{\# faces in } K)$$
$$(= K\text{:n kärkien lkm.} - K\text{:n sivujen lkm.} + K\text{:n tahkojen lkm.})$$

Example: $\chi(T^2) = 16 - 32 + 16 = 0.$
$\chi(S^2) = \chi(\text{cube}) = 8 - 12 + 6 = 2.$

4.2 Homotopy Groups

4.2.1 Paths and Loops

Let X be a topological space, $I = [0, 1] \subset \mathbb{R}$.
A continuous map $\alpha : I \to X$ is a **path** in X. The path α starts at $\alpha_0 = \alpha(0)$ and ends at $\alpha_1 = \alpha(1)$.
If $\alpha_0 = \alpha_1 \equiv x_0$, then α is a **loop** with **base point** x_0. We will focus on loops.

Definition: A **product** of two loops α, β with the same base point x_0, denoted by $\alpha * \beta$, is the loop

$$(\alpha * \beta)(t) = \begin{cases} \alpha(2t) & 0 \le t \le \frac{1}{2} \\ \beta(2t-1) & \frac{1}{2} \le t \le 1 \end{cases}$$

4.2.2 Homotopy

Let α, β be two loops in X with base point x_0. α and β are **homotopic**, $\alpha \sim \beta$, if there exists a continuous map $F : I \times I \to X$ such that

$$\begin{aligned} F(s,0) &= \alpha(s) && \forall s \in I \\ F(s,1) &= \beta(s) && \forall s \in I \\ F(0,t) &= F(1,t) = x_0 && \forall t \in I. \end{aligned}$$

F is called a **homotopy** between α and β.

Homotopy is an equivalence relation:

1. $\alpha \sim \alpha$: choose $F(s,t) = \alpha(s) \quad \forall t \in I$

2. $\alpha \sim \beta,$ homotopy $F(s,t) \Rightarrow \beta \sim \alpha,$ homotopy $F(s, 1-t)$

3. $\alpha \sim \beta,$ homotopy $F(s,t);$ $\beta \sim \gamma,$ homotopy $G(s,t)$. Then choose

$$H(s,t) = \begin{cases} F(s, 2t) & 0 \le t \le \frac{1}{2} \\ G(s, 2t-1) & \frac{1}{2} \le t \le 1 \end{cases}$$

$\Rightarrow$ $H(s,t)$ is a homotopy between α and γ, so $\alpha \sim \gamma$.

The equivalence class $[\alpha]$ is called the **homotopy class** of α.
($[\alpha] = \{$ all paths homotopic with α $\}$).

Lemma: If $\alpha \sim \alpha'$ and $\beta \sim \beta'$, then $\alpha * \beta \sim \alpha' * \beta'$.
Proof: Let $F(s,t)$ be a homotopy between α and α' and let $G(s,t)$ be a homotopy between β and β'. Then

$$H(s,t) = \begin{cases} F(2s, t) & 0 \le s \le \frac{1}{2} \\ G(2s-1, t) & \frac{1}{2} \le s \le 1 \end{cases}$$

is a homotopy between $\alpha * \beta$ and $\alpha' * \beta'$. $\square$

By the lemma, we can define a product of homotopy classes: $\quad [\alpha] * [\beta] \equiv [\alpha * \beta]$.

Theorem: The set of homotopy classes of loops at $x_0 \in X$, with the product defined as above, is a group called the **fundamental group** (or **first homotopy group**) of X at x_0. It is denoted by $\Pi_1(X, x_0)$
Proof:

(0) Closure under multiplication: For all $[\alpha], [\beta] \in \Pi_1(X, x_0)$ we have $[\alpha] * [\beta] = [\alpha * \beta] \in \Pi_1(X, x_0)$, since $\alpha * \beta$ is also a loop at x_0.

(1) Associativity: We need to show $(\alpha * \beta) * \gamma \sim \alpha * (\beta * \gamma)$.

$$\text{Homotopy } F(s,t) = \begin{cases} \alpha\left(\frac{4s}{1+t}\right) & 0 \le s \le \frac{1+t}{4} \\ \beta(4s - t - 1) & \frac{1+t}{4} \le s \le \frac{2+t}{4} \\ \gamma\left(\frac{4s-t-2}{2-t}\right) & \frac{2+t}{4} \le s \le 1 \end{cases}$$

$\Rightarrow$ $[(\alpha * \beta) * \gamma] = [\alpha * (\beta * \gamma)] \equiv [\alpha * \beta * \gamma]$.

(2) Unit element: Let us show that the unit element is $e = [C_{x_0}]$, where C_{x_0} is the constant path $C_{x_0}(s) = x_0 \quad \forall s \in I$. This follows since we have the homotopies:

$$\alpha * C_{x_0} \sim \alpha : \qquad F(s,t) = \begin{cases} \alpha\left(\frac{2s}{1+t}\right) & 0 \le s \le \frac{1+t}{2} \\ x_0 & \frac{1+t}{2} \le s \le 1 \end{cases}$$

$$C_{x_0} * \alpha \sim \alpha : \qquad F(s,t) = \begin{cases} x_0 & 0 \le s \le \frac{1-t}{2} \\ \alpha\left(\frac{2s-1+t}{1+t}\right) & \frac{1-t}{2} \le s \le 1 \end{cases} .$$

$\Rightarrow \; [\alpha * C_{x_0}] = [C_{x_0} * \alpha] = [\alpha]$.

(3) Inverse: Define $\alpha^{-1}(s) = \alpha(1-s)$. We need to show that α^{-1} is really the inverse of α: $[\alpha * \alpha^{-1}] = [C_{x_0}]$. Define:

$$F(s,t) = \begin{cases} \alpha(2s(1-t)) & 0 \le s \le \frac{1}{2} \\ \alpha(2(1-s)(1-t)) & \frac{1}{2} \le s \le 1 \end{cases}$$

Now we have $F(s,0) = \alpha * \alpha^{-1}$ and $F(s,1) = C_{x_0}$ so $\alpha * \alpha^{-1} \sim C_{x_0}$. Similarly $\alpha^{-1} * \alpha \sim C_{x_0}$ so we have proven the claim: $[\alpha^{-1} * \alpha] = [\alpha * \alpha^{-1}] = [C_{x_0}]$. $\square$

4.2.3 Properties of the Fundamental Group

1. If x_0 and x_1 can be connected by a path, then $\Pi_1(X, x_0) \cong \Pi_1(X, x_1)$. If X is arcwise connected, then the fundamental group is independent of the choice of x_0 up to an isomorphism: $\Pi_1(X, x_0) \cong \Pi_1(X)$.

   (A space X is arcwise connected if any two points $x_0, x_1 \in X$ can be connected with a path. It can be shown that an arcwise connected space is always connected, but the converse is not true.)

2. $\Pi_1(X)$ is a topological invariant: $X \approx Y \Rightarrow \Pi_1(X) \cong \Pi_1(Y)$.

3. Examples:

   - $\Pi_1(\mathbb{R}^2) = 0$ (= the trivial group)
   - $\Pi_1(T^2) = \Pi_1(S^1 \times S^1) = \mathbb{Z} \times \mathbb{Z}$.

   (One can show that $\Pi_1(X \times Y) = \Pi_1(X) \times \Pi_1(Y)$ for arcwise connected spaces X and Y.)

The real projective space is defined as $\mathbb{R}P^n = \{$ lines through the origin in $\mathbb{R}^{n+1}\}$. If $x = (x^0, x^1, \ldots, x^n) \neq 0$, then x defines a line. All $y = \lambda x$ for some nonzero $\lambda \in \mathbb{R}$ are on the same line and thus we have an equivalence relation: $y \sim x \Leftrightarrow y = \lambda x, \; \lambda \in \mathbb{R} - \{0\} \Leftrightarrow$ (x and y are on the same line.)
So $\mathbb{R}P^n = \{[x] | \; x \in \mathbb{R}^{n+1} - 0\}$ with the above equivalence relation.

Example: $\mathbb{R}P^2 \approx$ (S^2 with opposite points identified)
$\Pi_1(\mathbb{R}P^2) = \mathbb{Z}_2$.

4.2.4 Higher Homotopy Groups

Define: $I^n = \{(s_1, \ldots, s_n) | \ 0 \leq s_i \leq 1, \ 1 \leq i \leq n\}$

$\partial I^n =$ boundary of $I^n = \{(s_1, \ldots, s_n) | \text{ some } s_i = 0 \text{ or } 1\}$

A map $\alpha : I^n \to X$ which maps every point on ∂I^n to the same point $x_0 \in X$ is called an **n-loop** at $x_0 \in X$. Let α and β be n-loops at x_0. We say that α is homeotopic to β, $\alpha \sim \beta$, if there exists a continuous map $F : I^n \times I \to X$ such that

$$\begin{aligned} F(s_1, \ldots, s_n, 0) &= \alpha(s_1, \ldots, s_n) \\ F(s_1, \ldots, s_n, 1) &= \beta(s_1, \ldots, s_n) \\ F(s_1, \ldots, s_n, t) &= x_0 \quad \forall t \in I \text{ when } (s_1, \ldots, s_n) \in \partial I^n. \end{aligned}$$

Homotopy $\alpha \sim \beta$ is again an equivalence relation with respect to homotopy classes $[\alpha]$.

$$\text{Define: } \alpha * \beta : \ \alpha * \beta(s_1, \ldots, s_n) = \begin{cases} \alpha(2s_1, s_2, \ldots, s_n) & 0 \leq s_1 \leq \frac{1}{2} \\ \beta(2s_1 - 1, s_2, \ldots, s_n) & \frac{1}{2} \leq s_1 \leq 1. \end{cases}$$

$\alpha^{-1} : \ \alpha^{-1}(s_1, \ldots, s_n) = \alpha(1 - s_1, \ldots, s_n)$

$[\alpha] * [\beta] = [\alpha * \beta]$

$\Rightarrow \Pi_n(X, x_0)$, the $\mathbf{n^{th}}$ **homotopy group** of X at x_0. (This classifies continuous maps $S^n \to X$.)

Example: $\Pi_2(S^2) = \mathbb{Z}$.

4.3 Differentiable Manifolds

Definition: M is an **m-dimensional differentiable manifold** if

(i) M is a topological space

(ii) M is provided with a family of pairs $\{(U_i, \varphi_i)\}$, where $\{U_i\}$ is an open covering of M: $\bigcup_i U_i = M$, and every $\varphi_i : U_i \to U_i' \subset \mathbb{R}^m$, U_i' open, is a homeomorphism.

- The pair (U_i, φ_i) is called a **chart**, $\{(U_i, \varphi_i)\}$ an **atlas**, U_i the **coordinate neighbourhood** and φ_i the **coordinate function**.
  $\varphi(p) = (x^1(p), \ldots, x^m(p)), \ p \in U_i$ are the coordinate(s) of p.

(iii) Given U_i and U_j such that $U_i \bigcap U_j \neq \emptyset$, the map $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$ from $\varphi_j(U_i \bigcap U_j)$ to $\varphi_i(U_i \bigcap U_j)$ is **infinitely differentiable** (or: C^∞ or **smooth**).

- ψ_{ij} is called a **transition function**.

Recall: $f : \mathbb{R}^m \to \mathbb{R}^n$ is C^k if the partial derivatives

$$\frac{\partial^k f^l}{\partial(x^1)^{k_1} \cdots \partial(x^m)^{k_m}}, \qquad f = (f^1, \ldots, f^n), \qquad \begin{array}{l} l = 1, \ldots, n \\ k_1 + k_2 + \ldots + k_n = k \end{array}$$

exist and are continuous. The function f is C^∞ if all partial derivatives exist and are continuous for any k. We also call a C^∞ function f smooth.

The number m is the **dimension** of the manifold: dim $M = m$.

If the union of two atlases $\{(U_i, \varphi_i)\}$, $\{(V_i, \psi_i)\}$ is again an atlas, they are said to be **compatible**. This gives an equivalence relation among atlases, the equivalence class is called a **differentiable structure**.

A given differentiable manifold M can have several different differentiable structures: for example S^7 has 28 and $\mathbb{R}^4$ has infinitely (!) many differentiable structures.

Examples of differentiable manifolds: S^n
Let's realize S^n as a subset of $\mathbb{R}^{n+1}$: $S^n = \{x \in \mathbb{R}^{n+1} | \sum_{i=0}^n (x^i)^2 = 1\}$.
One possible atlas:

- coordinate neighbourhoods:

$$U_{i+} \equiv \{x \in S^n | x^i > 0\}$$
$$U_{i-} \equiv \{x \in S^n | x^i < 0\}$$

- coordinates:

$$\varphi_{i+}(x^0, \dots, x^n) = (x^0, \dots, x^{i-1}, x^{i+1}, \dots, x^n) \in \mathbb{R}^n$$
$$\varphi_{i-}(x^0, \dots, x^n) = (x^0, \dots, x^{i-1}, x^{i+1}, \dots, x^n) \in \mathbb{R}^n$$

  (so these are projections on the plane $x^i = 0$.)

The transition functions ($i \neq j,\ \alpha = \pm,\ \beta = \pm$),

$$\begin{aligned}\psi_{i\alpha j\beta} =& \varphi_{i\alpha} \circ \varphi_{j\beta}^{-1}, \\ (x^0, \dots, & x^i, \dots, x^{j-1}, x^{j+1}, \dots, x^n) \\ & \mapsto (x^0, \dots, x^{i-1}, x^{i+1}, \dots, x^{j-1}, \beta \sqrt{1 - \sum_{k \neq j} (x^k)^2}, x^{j+1}, \dots, x^n)\end{aligned}$$

are C^∞.
There are other compatible atlases, e.g. the stereographic projection.

4.3.1 Manifold with a Boundary

Let $\mathbb{H}$ be the "upper" half-space: $\mathbb{H}^m = \{(x^1, \dots, x^m) \in \mathbb{R}^m \mid x^m \geq 0\}$.
Now require for the coordinate functions: $\varphi_i : U_i \to U_i' \subset \mathbb{H}^m$, where U_i' is open in

$\mathbb{H}^m$. (The topology on $\mathbb{H}^m$ is the relative topology induced from $\mathbb{R}^m$.)
Points with coordinate $x^m = 0$ belong to the **boundary** of M (denoted by ∂M). The transition functions must now satisfy: $\psi_{ij} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$ are C^∞ in an open set of $\mathbb{R}^m$ which contains $\varphi_j(U_i \cap U_j)$. .

4.4 The Calculus on Manifolds

4.4.1 Differentiable Maps

Let M, N be differentiable manifolds with dimensions dim $M = m$ and dim $N = n$. Let f be a map $f : M \to N,\ p \mapsto f(p)$. Take charts (U, φ) and (V, ψ) such that $p \in U$ and $f(p) \in V$. If the combined map $\psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ is C^∞ at $\varphi(p)$, then f is **differentiable** at p. The definition is independent of the choice of charts, since if (U_1, φ_1) is some other chart at p, then

$$\psi \circ f \circ \varphi_1^{-1} = \overbrace{\psi \circ f \circ \varphi^{-1}}^{C^\infty} \circ \overbrace{\varphi \circ \varphi_1^{-1}}^{C^\infty} \qquad \Rightarrow \psi \circ f \circ \varphi_1^{-1} \text{ is } C^\infty.$$

If in addition $\psi \circ f \circ \varphi^{-1}$ is invertible, i.e. the inverse map $\varphi \circ f^{-1} \circ \psi^{-1}$ exists and is also C^∞, then f is called a **diffeomorphism** between M and N. In this case we say that M is diffeomorphic to N and denote it by $M \equiv N$.

Note:	homeomorphism	=	continuous deformation
	diffeomorphism	=	smooth deformation

- An **open curve** on M is a map $c :]a, b[\to M$ where $]a, b[$ is an open interval in $\mathbb{R}$ (notation: $(a, b) =]a, b[$).
- A **closed curve** is a map $S^1 \to M$.
- On a chart (U, φ) a curve c has a coordinate representation
  $x(t) = (\varphi \circ c)(t) : \mathbb{R} \to \mathbb{R}^m$.

A **function** f on M is a smooth map $M \to \mathbb{R}$.
$\mathcal{F} =$ the set of smooth maps $= \{f : M \to \mathbb{R} | f \text{ is smooth}\}$.

4.4.2 Tangent Vectors

Tangent vectors are defined using curves. Let $c : (a, b) \to M$ be a curve (we can assume $0 \in (a, b)$). Denote $c(0) = p$ and let $f : M \to \mathbb{R}$ be a function.
The rate of change of f along the curve c at point p is

$$\left.\frac{df(c(t))}{dt}\right|_{t=0} = \left.\frac{\partial f}{\partial x^\mu}\frac{dx^\mu(c(t))}{dt}\right|_{t=0},$$

where $x^\mu(p) = \varphi^\mu(p)$ are local coordinates and

$$\frac{\partial f}{\partial x^\mu} \equiv \frac{\partial(f \circ \varphi^{-1}(x))}{\partial x^\mu}.$$

Also we have introduced the Einstein summation convention:

- When an index appears once as a subscript and once as a superscript, it is understood to be summed over. For example $x_\mu y^\mu \equiv \sum_{\mu=1}^m x_\mu y^\mu = x_1 y^1 + \ldots + x_m y^m$.

In other words, $\frac{df(c(t))}{dt}$ is obtained by acting on the function f with the differential operator

$$X_p \equiv X_p^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p, \text{ where } X_p^\mu = \left.\frac{dx^\mu(c(t))}{dt}\right|_{t=0}.$$

The operator X_p is called a **tangent vector** of M at p. It depends on the curve, but several curves can give rise to the same tangent vector X_p. We can see that two curves c_1 and c_2 give the same X_p if and only if

(i) $c_1(0) = c_2(0) = p$

(ii) $\left.\frac{dx^\mu(c_1(t))}{dt}\right|_{t=0} = \left.\frac{dx^\mu(c_2(t))}{dt}\right|_{t=0}$

This gives an equivalence relation between the two curves, $c_1 \sim c_2$. Thus equivalence classes can be identified with tangent vectors X_p.

The set of all tangent vectors at p is the **tangent space** T_pM at p. It is a real vector space, dim $T_pM = m$:

- $X_{1p} + X_{2p} = (X_{1p}^\mu + X_{2p}^\mu)\left(\frac{\partial}{\partial x^\mu}\right)_p$
- $cX_p = (cX_p^\mu)\left(\frac{\partial}{\partial x^\mu}\right)_p$

$(e_\mu)_p = \left(\frac{\partial}{\partial x^\mu}\right)_p$ is called the **coordinate basis**.

The vectors are independent of a choice of coordinates, if their components are transformed in a correct way. Let $x(p) = \varphi_i(p)$ and $y(p) = \varphi_j(p)$ be two coordinates. For the vector to be independent of the choice of coordinates we must have

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

But on the other hand by the chain rule we have

$$X^\mu \frac{\partial}{\partial x^\mu} = X^\nu \frac{\partial y^\mu}{\partial x^\nu} \frac{\partial}{\partial y^\mu}.$$

Thus we get the transformation rule for the components:

$$\boxed{Y^\mu = X^\nu \frac{\partial y^\mu}{\partial x^\nu}}$$

Note the abuse of the notation:

$$X^\nu \frac{\partial y^\mu}{\partial x^\nu} \frac{\partial}{\partial x^\mu} \equiv X^\nu_p \frac{\partial(\varphi_j \circ \varphi_i^{-1})(x^\mu(p))}{\partial x^\nu(p)} \left(\frac{\partial}{\partial x^\mu} \right)_p .$$

Let us now leave calculus on manifolds for a while and study vector spaces some more.

4.4.3 Dual Vector Space

Let V be a complex vector space and f a linear function $V \to \mathbb{C}$. Now $V^* = \{f | f \text{ is a linear function } V \to \mathbb{C}\}$ is also a complex vector space, the **dual vector space** to V:

- $(f_1 + f_2)(\vec{v}) = f_1(\vec{v}) + f_2(\vec{v})$
- $(af)(\vec{v}) = a(f(\vec{v}))$
- $\vec{0}_{V^*}(\vec{v}) = 0 \quad \forall \vec{v} \in V$

The elements of V^* are called the **dual vectors**.
Let $\{\vec{e}_1, \ldots, \vec{e}_n\}$ be a basis of V. Then any vector $\vec{v} \in V$ can be written as $\vec{v} = v^i \vec{e}_i$. We define a **dual basis** in V^* such that $e^{*i}(\vec{e}_j) = \delta^i{}_j$. From this it follows that $\dim V = \dim V^* = n$ (dual basis $= \{e^{*1}, \ldots, e^{*n}\}$). We can then expand any $f \in V^*$ as $f = f_i e^{*i}$ for some coefficients $f_i \in \mathbb{C}$. Now we have

$$f(\vec{v}) = f_i e^{*i}(v^j \vec{e}_j) = f_i v^j e^{*i}(\vec{e}_j) = f_i v^i.$$

This can be interpreted as an **inner product**:

$$\begin{aligned} < \, , \, > : \ & V^* \times V \to \mathbb{C} \\ < f, \vec{v} > = \ & f_i v^i. \end{aligned}$$

(Note that this is not the same inner product $< \, | \, >$ which we discussed before: $< , > : \ V^* \times V \to \mathbb{C}$ but $< \ | \ > : \ V \times V \to \mathbb{C}$.)

Pullback: Let $f : V \to W$ and $g : W \to \mathbb{C}$ be linear maps ($g \in W^*$). It follows that $g \circ f : V \to \mathbb{C}$ is a linear map, i.e. $g \circ f \in V^*$.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ & \searrow & \downarrow g \\ g \circ f & & \mathbb{C} \end{array}$$

Now f induces a map $f^* : W^* \to V^*, g \mapsto g \circ f$ i.e. $f^*(g) = g \circ f \in V^*$. $f^*(g)$ is called the **pullback** (takaisinveto) of g.

Dual of a Dual: Let $\omega : V^* \to \mathbb{C}$ be a linear function ($\omega \in (V^*)^*$). Every $\vec{v} \in V$ induces via the inner product a mapping $\omega_{\vec{v}} \in (V^*)^*$ defined by $\omega_{\vec{v}}(f) =< f, \vec{v} >$. On the other hand, it can be shown this gives all $\omega \in (V^*)^*$. So we can identify $(V^*)^*$ with V.

Tensors: A tensor of type (p, q) is a function of p dual vectors and q vectors, and is linear in its every argument[1]

$$T : \overbrace{V^* \times \ldots \times V^*}^{p} \times \overbrace{V \times \ldots \times V}^{q} \to \mathbb{C}.$$

Examples: (0,1) tensor = dual vector : $V \to \mathbb{C}$
(1,0) tensor = (dual of a dual) vector
(1,2) tensor: $T : V^* \times V \times V \to \mathbb{C}$. Choose basis $\{\vec{e}_i\}$ in V and $\{e^{*i}\}$ in V^*:

$$T(f, \vec{v}, \vec{w}) = T(f_i e^{*i}, v^j \vec{e}_j, w^k \vec{e}_k) = f_i v^j w^k \overbrace{T(e^{*i}, \vec{e}_j, \vec{e}_k)}^{\equiv T^i{}_{jk}} = T^i{}_{jk} f_i v^j w^k,$$

where $T^i{}_{jk}$ are the components of the tensor and they uniquely determine the tensor. Note the positioning of the indices.
In general, (p, q) tensor components have p upper and q lower indices.

Tensor product: Let R be a (p, q) tensor and S be a (p', q') tensor. Then $T = R \otimes S$ is defined as the $(p + p', q + q')$ tensor:

$$\begin{aligned} &T(f_1, \ldots, f_p; f_{p+1}, \ldots, f_{p+p'}; \vec{v}_1, \ldots, \vec{v}_q; \vec{v}_{q+1}, \ldots, \vec{v}_{q+q'}) \\ &\qquad = R(f_1, \ldots, f_p; \vec{v}_1, \ldots, \vec{v}_q) S(f_{p+1}, \ldots, f_{p+p'}; \vec{v}_{q+1}, \ldots, \vec{v}_{q+q'}). \end{aligned}$$

In terms of components:

$$T^{i_1 \ldots i_p i_{p+1} \ldots i_{p+p'}}_{j_1 \ldots j_q j_{q+1} \ldots j_{q+q'}} = R^{i_1 \ldots i_p}_{j_1 \ldots j_q} S^{i_{p+1} \ldots i_{p+p'}}_{j_{q+1} \ldots j_{q+q'}}$$

Contraction: This is an operation that produces a $(p-1, q-1)$ tensor from a (p, q) tensor:

$$\underbrace{T}_{(p,q)} \mapsto \underbrace{T_{c(ij)}}_{(p-1,q-1)} ,$$

where the $(p - 1, q - 1)$ tensor $T_{c(ij)}$ is

$$T_{c(ij)}(f_1, \ldots, f_{p-1}; \vec{v}_1, \ldots, \vec{v}_{q-1}) = T(f_1, \ldots, \overbrace{e^{*k}}^{i^{th}}, \ldots, f_{p-1}; \vec{v}_1, \ldots, \overbrace{\vec{e}_k}^{j^{th}}, \ldots, \vec{v}_{q-1}).$$

Note the sum over k in the formula above. In component form this is

$$T_{c(ij)}{}^{l_1 \ldots l_{p-1}}_{m_1 \ldots m_{q-1}} = T^{l_1 \ldots l_{i-1} k l_i \ldots l_{p-1}}_{m_1 \ldots m_{j-1} k m_j \ldots m_{q-1}}$$

Now we can return to calculus on manifolds.

[1]So T is a multilinear object.

4.4.4 1-forms (i.e. cotangent vectors)

Tangent vectors of a differentiable manifold M at point p were elements of the vector space T_pM. **Cotangent vectors** or **1-forms** are their dual vectors, i.e. linear functions $T_pM \to \mathbb{R}$. In other words, they are elements of the dual vector space T_p^*M. Let $w \in T_p^*M$ and $v \in T_pM$, then the inner product $< , >: T_p^*M \times T_pM \to \mathbb{R}$ is

$$< w, v >= w(v) \in \mathbb{R}.$$

The inner product is bilinear:

$$\begin{aligned} < w, \alpha_1 v_1 + \alpha_2 v_2 > &= w(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 < w, v_1 > +\alpha_2 < w, v_2 > \\ < \alpha_1 w_1 + \alpha_2 w_2, v > &= (\alpha_1 w_1 + \alpha_2 w_2)(v) = \alpha_1 < w_1, v > +\alpha_2 < w_2, v > . \end{aligned}$$

Let $\{e_\mu\} = \{\frac{\partial}{\partial x^\mu}\}$ be a coordinate basis of T_pM. (Note that the correct notation would be $\{\left(\frac{\partial}{\partial x^\mu}\right)_p\}$, but this is somewhat cumbersome so we use the shorter notation.) The dual basis is denoted by $\{dx^\mu\}$ and it satisfies by definition

$$< dx^\mu, \frac{\partial}{\partial x^\nu} >= dx^\mu(\frac{\partial}{\partial x^\nu}) = \delta^\mu{}_\nu.$$

Now we can expand $w = w_\mu dx^\mu$ and $v = v^\nu \frac{\partial}{\partial x^\nu}$. Then

$$w(v) =< w, v >= w_\mu v^\nu dx^\mu(\frac{\partial}{\partial x^\nu}) = w_\mu v^\mu.$$

Consider now a function $f \in \mathcal{F}(M)$ (i.e. f is a smooth map $M \to \mathbb{R}$). Its **differential** $df \in T_p^*M$ is the map

$$df(v) =< df, v >\equiv v(f) = v^\mu \frac{\partial f}{\partial x^\mu}.$$

Thus the components of df are $\frac{\partial f}{\partial x^\mu}$ and

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu.$$

Consider two coordinate patches U_i and U_j with $p \in U_i \cap U_j$. Let $x = \varphi_i(p)$ and $y = \varphi_j(p)$ be the coordinates in U_i and U_j respectively. We can derive how the components of a 1-form transform under the change of coordinates:
Let $w = w_\mu dx^\mu = \tilde{w}_\nu dy^\nu \in T_p^*M$ and $v = v^\rho \frac{\partial}{\partial x^\rho} = \tilde{v}^\sigma \frac{\partial}{\partial y^\sigma} \in T_pM$ be a 1-form and a vector. We already know that $\tilde{v}^\nu = \frac{\partial y^\nu}{\partial x^\mu} v^\mu$, so we get

$$w(v) = w_\mu v^\mu = \tilde{w}_\nu \tilde{v}^\nu = \tilde{w}_\nu \frac{\partial y^\nu}{\partial x^\mu} v^\mu,$$

so we find the transformed components

$$w_\mu = \tilde{w}_\nu \frac{\partial y^\nu}{\partial x^\mu} \qquad \text{or} \quad \tilde{w}_\mu = w_\nu \frac{\partial x^\nu}{\partial y^\mu}.$$

The dual basis vectors transform as

$$dy^\nu = \frac{\partial y^\nu}{\partial x^\mu} dx^\mu.$$

4.4.5 Tensors on a manifold

A tensor of type (q,r) is a multilinear map

$$T:\overbrace{T_p^*M\times\ldots\times T_p^*M}^{q}\times\overbrace{T_pM\times\ldots\times T_pM}^{r}\to\mathbb{R}.$$

Denote the set of type (q,r) tensors at $p\in M$ by $T^q_{r,p}(M)$. Note that $T^1_{0,p}=(T_p^*M)^*=T_pM$ and $T^0_{1,p}(M)=T_p^*M$.
The basis of $T^q_{r,p}$ is

$$\left\{\frac{\partial}{\partial x^{\mu_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{\mu_q}}\otimes dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_r}\right\}.$$

The basis vectors satisfy (as a mapping $T_p^*M\times\ldots\times T_p^*M\times T_pM\times\ldots\times T_pM\to\mathbb{R}$):

$$\left(\frac{\partial}{\partial x^{\mu_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{\mu_q}}\otimes dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_r}\right)\left(dx^{\alpha_1},\ldots,dx^{\alpha_q},\frac{\partial}{\partial x^{\beta_1}},\ldots,\frac{\partial}{\partial x^{\beta_r}}\right)$$
$$=\delta^{\alpha_1}{}_{\mu_1}\ldots\delta^{\alpha_q}{}_{\mu_q}\delta^{\nu_1}{}_{\beta_1}\ldots\delta^{\nu_r}{}_{\beta_r}.$$

(Note that $\frac{\partial}{\partial x^\mu}(dx^\alpha)\equiv< dx^\alpha,\frac{\partial}{\partial x^\mu}>=\delta^\alpha{}_\mu$. On the left $\frac{\partial}{\partial x^\mu}$ is interpreted as an element of $(T_p^*M)^*$.)
We can expand as $T=T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\nu_r}\left\{\frac{\partial}{\partial x^{\mu_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{\mu_q}}\otimes dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_r}\right\}$ so

$$T(w_1,\ldots,w_q;v_1,\ldots,v_r)=T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\nu_r}w_{1\mu_1}\ldots w_{q\mu_q}v_1^{\nu_1}\ldots v_r^{\nu_r}.$$

The **tensor product** of tensors $T\in T^q_{r,p}(M)$ and $U\in T^s_{t,p}(M)$ is the tensor $T\otimes U\in T^{q+s}_{r+t,p}(M)$ with

$$\begin{aligned}(T\otimes U)&(w_1,\ldots,w_q,w_{q+1},\ldots,w_{q+s};v_1,\ldots,v_r,v_{r+1},\ldots,v_{r+t})\\&=T(w_1,\ldots,w_q;v_1,\ldots,v_r)U(w_{q+1},\ldots,w_{q+s};v_{r+1},\ldots,v_{r+t}).\\&=T^{\mu_1\ldots\mu_q}{}_{\nu_1\ldots\nu_r}w_{1\mu_1}\ldots w_{q\mu_q}v_1^{\nu_1}\ldots v_r^{\nu_r}\cdot\\&\qquad U^{\alpha_1\ldots\alpha_s}{}_{\beta_1\ldots\beta_t}w_{(q+1)\alpha_1}\ldots w_{(q+s)\alpha_s}v_{r+1}^{\beta_1}\ldots v_{r+t}^{\beta_t}.\end{aligned}$$

Contraction maps a tensor $T\in T^q_{r,p}(M)$ to a tensor $T'\in T^{q-1}_{r-1,p}(M)$ with components

$$T'^{\mu_1\ldots\mu_{q-1}}{}_{\nu_1\ldots\nu_{r-1}}=T^{\mu_1\ldots\mu_{i-1}\rho\mu_i\ldots\mu_{q-1}}{}_{\nu_1\ldots\nu_{j-1}\rho\nu_j\ldots\nu_{r-1}}$$

Under a coordinate transformation, a tensor of type (q,r) transforms like a product of q vectors and r one-forms (note that $v_1\otimes\cdots\otimes v_q\otimes w_1\otimes\ldots\otimes w_r$ is one example of a (q,r) tensor). For example $T\in T^1_{2,p}(M)$ tensor of type $(1,2)$:

$$T=T^\alpha{}_{\beta_1\beta_2}\frac{\partial}{\partial x^\alpha}\otimes dx^{\beta_1}\otimes dx^{\beta_2}=\tilde{T}^\mu{}_{\nu_1\nu_2}\frac{\partial}{\partial y^\mu}\otimes dy^{\nu_1}\otimes dy^{\nu_2}$$

gives us the transformation rule for the components

$$\tilde{T}^\mu{}_{\nu_1\nu_2}=\frac{\partial y^\mu}{\partial x^\alpha}\frac{\partial x^{\beta_1}}{\partial y^{\nu_1}}\frac{\partial x^{\beta_2}}{\partial y^{\nu_2}}T^\alpha{}_{\beta_1\beta_2}$$

4.4.6 Tensor Fields

Suppose that a vector $v(p)$ has been assigned to every point p in M. This is a **(smooth) vector field**, if for every C^∞ function $f \in \mathcal{F}$ the function $v(p)(f) : M \to \mathbb{R}$ is also a smooth function. We denote $v(p)(f)$ by $v[f]$. The set of smooth vector fields on M is denoted by $\chi(M)$.

Smooth cotangent vector field : For every $p \in M$ there is $w(p) \in T_p^*M$ such that if $V \in \chi(M)$, then the function

$$\begin{aligned} w[V] \quad &: \quad M \to \mathbb{R} \\ p \quad &\mapsto \quad w[V](p) = w(p)(V(p)) \end{aligned}$$

is smooth. The set of cotangent vector fields is denoted by $\Omega^1(M)$.

Smooth (q,r)-tensor field : If for all $p \in M$ there is $T(p) \in T^q_{r,p}(M)$ such that if $w_1, \ldots, w_q$ are smooth cotangent vector fields and $v_1, \ldots, v_r$ are smooth tangent vector fields, then the map

$$p \mapsto T[w_1, \ldots, w_q; v_1, \ldots, v_r](p) = T(p)(w_1(p), \ldots, w_q(p); v_1(p), \ldots, v_r(p))$$

is smooth on M.

4.4.7 Differential Map and Pullback

Let M and N be differentiable manifolds and $f : M \to N$ smooth.
f induces a map called the **differential map** (työntökuvaus) $f_* : T_pM \to T_{f(p)}N$. It is defined as follows:
If $g \in \mathcal{F}(N)$ (i.e. $g : N \to \mathbb{R}$ smooth), and $v \in T_pM$, then

$$(f_*v)[g] = v[g \circ f].$$

In other words, if v characterizes the rate of change of a function along a curve $c(t)$, then f_*v characterizes the rate of change of a function along the curve $f(c(t))$.
Let x be local coordinates on M and y be local coordinates on N, "$y = f(x)$". Also let $v = v^\mu \frac{\partial}{\partial x^\mu}$ and $f_*v = (f_*v)^\nu \frac{\partial}{\partial y^\nu}$. Then

$$v[g \circ f] = v^\mu \frac{\partial(g(f(x)))}{\partial x^\mu} = v^\mu \frac{\partial g}{\partial y^\nu} \frac{\partial y^\nu}{\partial x^\mu} \equiv (f_*v)^\nu \frac{\partial g}{\partial y^\nu}$$

and we get

$$(f_*v)^\nu = v^\mu \frac{\partial y^\nu}{\partial x^\mu}, \text{ where } y = f(x).$$

[More precisely $x^\mu = \varphi^\mu(p)$, $y^\nu = \psi^\nu(f(p))$ and $\frac{\partial y^\nu}{\partial x^\mu} = \frac{\partial(\psi \circ f \circ \varphi^{-1})^\nu}{\partial x^\mu}$.]

The function f also induces the map

$$f^* : T^*_{f(p)}N \to T^*_p M, \quad (f^* w)(v) = w(f_* v),$$

where $v \in T_p M$ and $w \in T^*_{f(p)}N$ are arbitrary. f^* is called the **pullback**.
In local coordinates, $w = w_\nu dy^\nu$,

$$w(f_* v) = w_\nu dy^\nu \left(v^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial}{\partial y^\alpha} \right) = w_\nu v^\mu \frac{\partial y^\nu}{\partial x^\mu} = (f^* w)_\mu v^\mu = (f^* w)(v),$$

from which we get

$$(f^* w)_\mu = w_\nu \frac{\partial y^\nu}{\partial x^\mu}.$$

The pullback f^* can also be generalized to $(0, r)$ tensors and similarly the differential map f_* can be generalized to $(q, 0)$ tensors.

4.4.8 Flow Generated by a Vector Field

Let X be a vector field on M. An **integral curve** $x(t)$ of X is a curve on M, whose tangent vector at $x(t)$ is $X|_{x(t)}$.
In local coordinates, the integral curve is the solution of the differential equations

$$\frac{dx^\mu(t)}{dt} = X^\mu(x(t)) \qquad \left(X = X^\mu \frac{\partial}{\partial x^\mu} \right).$$

The existence and uniqueness theorem of ordinary differential equations guarantees that the equation has a unique solution (at least locally in some neighbourhood of $t = 0$), once the initial condition $x^\mu(t = 0) = x^\mu_0$ has been specified. If M is compact, the solution exists for all t.
Let us denote the integral curve of X which passes the point x_0 at $t = 0$ by $\sigma(t, x_0)$. Thus

$$\begin{cases} \frac{d\sigma^\mu(t,x_0)}{dt} = X^\mu(\sigma(t, x_0)) \\ \sigma^\mu(t = 0, x_0) = x^\mu_0 \end{cases}.$$

The map $\sigma : I \times M \to M$ is called a **flow** generated by X ($I \subset \mathbb{R}$). It satisfies $\sigma(t, \sigma(s, x_0)) = \sigma(t + s, x_0)$ (as long as $t + s \in I$).
Proof: The left and right hand sides satisfy the same differential equation: $\frac{d}{dt}\sigma^\mu(t, \sigma) = X^\mu(\sigma) = \frac{d}{dt}\sigma^\mu(t + s, \sigma)$ and the same initial condition. Thus by uniqueness they are the same map. $\square$ (See Nakahara page 15)

For a fixed t, $\sigma(t, x)$ is a diffeomorphism $\sigma_t : M \to M,\ x \mapsto \sigma(t, x)$. The family of diffeomorphisms $\{\sigma_t | t \in I\}$ is a commutative (Abelian) group (when $I = \mathbb{R}$):

$$\begin{aligned} &\sigma_t \cdot \sigma_s \equiv \sigma_t \circ \sigma_s = \sigma_{t+s} \\ &\sigma_{-t} = (\sigma_t)^{-1} \\ &\sigma_0 = id_M. \end{aligned}$$

The group is called the **one-parameter group of transformations.**
Let $t = \epsilon$ be infinitesimally close to 0. Now,

$$\sigma^\mu_\epsilon(x) = \sigma^\mu(\epsilon, x) \approx \sigma^\mu(0, x) + \left.\frac{d\sigma^\mu(t,x)}{dt}\right|_{t=0} \epsilon + O(\epsilon^2) = x^\mu + X^\mu(x)\epsilon.$$

In this context the vector field X is called the **infinitesimal generator** of the transformation σ_t.

Given a vector field X, the corresponding flow is often denoted by

$$\sigma^\mu_t(x) = \sigma^\mu(t, x) = \exp(tX)x^\mu = (e^{tX})x^\mu$$

and called the exponentiation of X. This is because

$$\begin{aligned}
\sigma^\mu_t(x) &= x^\mu + t\left.\frac{d\sigma^\mu(s,x)}{ds}\right|_{s=0} + \frac{1}{2!}t^2 \left.\frac{d^2\sigma^\mu(s,x)}{ds^2}\right|_{s=0} + \cdots \\
&= \left(1 + t\frac{d}{ds} + \frac{1}{2!}t^2\frac{d^2}{ds^2} + \cdots\right) \sigma^\mu(s,x)\Big|_{s=0} \\
&= e^{t\frac{d}{ds}}\sigma^\mu(s,x)\Big|_{s=0} = e^{tX}x^\mu \ .
\end{aligned}$$

4.4.9 Lie Derivative

Let $\sigma_t(x)$ be a flow on M generated by vector field X: $\frac{d\sigma^\mu_t(x)}{dt} = X^\mu(\sigma_t(x))$. Let Y be another vector field on M. We want to calculate the rate of change of Y along the curve $x^\mu(t) = \sigma^\mu_t(x)$.
The **Lie derivative** of a vector field Y is defined by

$$\mathcal{L}_X Y = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left((\sigma_{-\epsilon})_* Y|_{\sigma_\epsilon(x)} - Y|_x\right).$$

Let's rewrite this in a more user-friendly form: First

$$\begin{aligned}
Y|_x &= Y^\mu(x)\frac{\partial}{\partial x^\mu} \\
Y|_{\bar{x}} &= Y^\mu(\bar{x})\frac{\partial}{\partial \bar{x}^\mu},
\end{aligned}$$

where we have for the coordinates

$$\begin{aligned}
\bar{x}^\mu \equiv& \sigma^\mu_\epsilon(x) = x^\mu + \epsilon X^\mu(x) + O(\epsilon^2) \\
&\Rightarrow x^\mu = \bar{x}^\mu - \epsilon X^\mu(\bar{x}) + O(\epsilon^2).
\end{aligned}$$

Thus

$$Y|_{\bar{x}} = (Y^\mu(x + \epsilon X))\frac{\partial}{\partial \bar{x}^\mu} = \left(Y^\mu(x) + \epsilon X^\nu \frac{\partial Y^\mu(x)}{\partial x^\nu}\right)\frac{\partial}{\partial \bar{x}^\mu}.$$

Differential map from $\bar{x}$ to x:

$$\begin{aligned}((\sigma_{-\epsilon})_* Y|_{\bar{x}})^\alpha = Y^\mu|_{\bar{x}} \frac{\partial x^\alpha}{\partial \bar{x}^\mu} &= \left(Y^\mu(x) + \epsilon X^\nu(x) \frac{\partial Y^\mu(x)}{\partial x^\nu}\right) \Bigg(\delta^\alpha{}_\mu - \epsilon \overbrace{\frac{\partial X^\alpha(\bar{x})}{\partial \bar{x}^\mu}}^{\frac{\partial X^\alpha}{\partial x^\mu} + O(\epsilon)}\Bigg) \\ &= Y^\alpha(x) + \epsilon \left(X^\nu(x) \frac{\partial Y^\alpha}{\partial x^\nu} - Y^\mu(x) \frac{\partial X^\alpha}{\partial x^\mu}\right) + O(\epsilon^2) \\ &\Rightarrow \mathcal{L}_X Y = \left(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\right) \frac{\partial}{\partial x^\mu}.\end{aligned}$$

So we got

$$\mathcal{L}_X Y = \left(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\right) \frac{\partial}{\partial x^\mu} = [X, Y],$$

where the commutator ("Lie bracket") acts on functions by

$$[X, Y]\ f = X[Y[f]] - Y[X[f]].$$

Note that XY is not a vector field but $[X, Y]$ is:

$$XY\ f = X[Y[f]] = X^\mu \partial_\mu [Y^\nu \partial_\nu f] = \underbrace{X^\mu(\partial_\mu Y^\nu)\partial_\nu}_{\text{vector field}} f + \underbrace{X^\mu Y^\nu \partial_\mu \partial_\nu}_{\text{not a vector field}} f.$$

Lie derivative of a one-form: Let $w \in \Omega^1(M)$ be a one-form (cotangent vector). Define the Lie derivative of w along X as

$$\mathcal{L}_X w = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\sigma_\epsilon^* w|_{\sigma_\epsilon(x)} - w|_x\right).$$

Let's simplify this. The coordinates at $\sigma_\epsilon(x) : y^\mu \equiv \sigma_\epsilon^\mu(x) \approx x^\mu + \epsilon X^\mu(x)$.

$$\begin{aligned}(\sigma_\epsilon^* w)_\alpha = w_\beta(y) \frac{\partial y^\beta}{\partial x^\alpha} &= w_\beta(x + \epsilon X) \frac{\partial}{\partial x^\alpha}(x^\beta + \epsilon X^\beta) \\ &= (w_\beta(x) + \epsilon X^\mu \partial_\mu w_\beta(x))(\delta^\beta{}_\alpha + \epsilon \partial_\alpha X^\beta) \\ &= w_\alpha + \epsilon(X^\mu \partial_\mu w_\alpha + w_\mu \partial_\alpha X^\mu)\end{aligned}$$

Thus we find

$$\mathcal{L}_X w = (X^\mu \partial_\mu w_\alpha + w_\mu \partial_\alpha X^\mu)\, dx^\alpha.$$

Lie derivative of a function: A natural guess would be $\mathcal{L}_X f = X[f]$. Let's check if this works:

$$\mathcal{L}_X f = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(f(\sigma_\epsilon(x)) - f(x)) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(f(x + \epsilon X) - f(x)) = X^\mu \partial_\mu f = Xf = X[f].$$

Thus the definition works.

Lie derivative of a tensor field: We define these using the Leibnitz rule: we require that

$$\mathcal{L}_X(t_1 \otimes t_2) = (\mathcal{L}_X t_1) \otimes t_2 + t_1 \otimes (\mathcal{L}_X t_2).$$

This is true if t_1 is a function ((0,0) tensor) and t_2 is a one form or a vector field, or vice versa. (exercise)
Example: Let's find the Lie derivative of a (1,1) tensor: $t = t_\mu{}^\nu dx^\mu \otimes e_\nu$; $e_\nu = \frac{\partial}{\partial x^\nu}$.

$$\begin{aligned}
\mathcal{L}_X t &= (\mathcal{L}_X t_\mu{}^\nu) dx^\mu \otimes e_\nu + t_\mu{}^\nu (\mathcal{L}_X dx^\mu) \otimes e_\nu + t_\mu{}^\nu dx^\mu \otimes (\mathcal{L}_X e_\nu) \\
&= (X^\alpha \partial_\alpha t_\mu{}^\nu) dx^\mu \otimes e_\nu + t_\mu{}^\nu (\partial_\alpha X^\mu) dx^\alpha \otimes e_\nu - t_\mu{}^\nu dx^\mu \otimes (\partial_\nu X^\alpha) e_\alpha \\
&= (X^\alpha \partial_\alpha t_\mu{}^\nu + t_\alpha{}^\nu \partial_\mu X^\alpha - t_\mu{}^\alpha \partial_\alpha X^\nu) dx^\mu \otimes e_\nu.
\end{aligned}$$

[We used here $e_\nu = \frac{\partial}{\partial x^\nu}$, $(e_\nu)^\alpha = \delta_\nu{}^\alpha$, $(dx^\mu)_\alpha = \delta^\mu{}_\alpha$, $(\mathcal{L}_X e_\nu)^\alpha = X^\mu \partial_\mu (e_\nu)^\alpha - (e_\nu)^\mu \partial_\mu X^\alpha = -\partial_\nu X^\alpha$ and also $(\mathcal{L}_X dx^\mu)_\alpha = X^\nu \partial_\nu (dx^\mu)_\alpha + (dx^\mu)_\nu \partial_\alpha X^\nu = \partial_\alpha X^\mu$.]

4.4.10 Differential Forms

A **differential form** of order r (or **r-form**) is a totally antisymmetric $(0, r)$-tensor:

$$p \in S_r : \; w(v_{p(1)}, \ldots, v_{p(r)}) = \text{sgn}(p) \; w(v_1, \ldots, v_r),$$

where sgn(p) is the sign of the permutation p:

$$\text{sgn}(p) = (-1)^{\text{number of exchanges}} = \begin{cases} +1 & \text{for an even permutation} \\ -1 & \text{for an odd permutation.} \end{cases}$$

Example: $p : (123) \to (231)$: Two exchanges $[(231) \to (213) \to (123)]$ to (123), thus p is an even permutation.
$\tilde{p} : (123) \to (321)$: One exchange to (231) and then two exchanges to (123), thus $\tilde{p}$ is an odd permutation.

The r-forms at point $p \in M$ form a vector space $\Omega^r_p(M)$. What is its basis?
We define the **wedge product** of 1-forms:

$$dx^{\mu_1} \wedge dx^{\mu_2} \wedge \ldots \wedge dx^{\mu_r} = \sum_{p \in S_r} \text{sgn}(p) \; dx^{\mu_{p(1)}} \otimes \ldots \otimes dx^{\mu_{p(r)}}$$

Then $\{ \; dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r} | \; \mu_1 < \mu_2 < \ldots < \mu_r \; \}$ forms the basis of $\Omega^r_p(M)$.

Examples: $dx^\mu \wedge dx^\nu = dx^\mu \otimes dx^\nu - dx^\mu \otimes dx^\nu$

$$\begin{aligned}
dx^1 \wedge dx^2 \wedge dx^3 = \; & dx^1 \otimes dx^2 \otimes dx^3 + dx^2 \otimes dx^3 \otimes dx^1 + dx^3 \otimes dx^1 \otimes dx^2 \\
& - dx^2 \otimes dx^1 \otimes dx^3 - dx^3 \otimes dx^2 \otimes dx^1 - dx^1 \otimes dx^3 \otimes dx^2.
\end{aligned}$$

Note:

- $dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r} = 0$ if the <u>same</u> index appears twice (or more times).
- $dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r} = \text{sgn}(p) dx^{\mu_{p(1)}} \wedge \ldots \wedge dx^{\mu_{p(r)}}$. (reshuffling of terms.)

In the above basis, an r-form $w \in \Omega_p^r(M)$ is expanded

$$w = \frac{1}{r!} w_{\mu_1 \ldots \mu_r} dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}.$$

Note: the components $w_{\mu_1 \ldots \mu_r}$ are totally antisymmetric in the indices
(e.g. $w_{\mu_1 \mu_2 \mu_3 \ldots \mu_r} = -w_{\mu_2 \mu_1 \mu_3 \ldots \mu_r}$).

One can show that dim $\Omega_p^r(M) = \frac{m!}{r!(m-r)!} = \binom{m}{r}$, where $m = \text{dim} M$.
Note also: $\Omega_p^1(M) = T_p^*(M)$ cotangent space
$\Omega_p^0(M) = \mathbb{R}$ by convention

Now we generalize the wedge product for the products of a q-form and an r-form and call it **exterior product**:

Definition: The exterior product of a q-form ω and an r-form η is a $(q+r)$-form $\omega \wedge \eta$:

$$(\omega \wedge \eta)(v_1, \ldots, v_{q+r}) = \frac{1}{q!r!} \sum_{p \in S_{q+r}} \text{sgn}(p) \omega(v_{p(1)}, \ldots, v_{p(q)}) \cdot \eta(v_{p(q+1)}, \ldots, v_{p(q+r)}).$$

If $q + r > m = \dim(M)$, then $\omega \wedge \eta = 0$. The exterior product satisfies the properties:

(i) $\omega \wedge \omega = 0$, if q is odd.

(ii) $\omega \wedge \eta = (-1)^{qr} \eta \wedge \omega$.

(iii) $(\omega \wedge \eta) \wedge \xi = \omega \wedge (\eta \wedge \xi)$.

[Proof: exercise]
We may assign an r-form smoothly at each point p on a manifold M, to obtain an r-form field. The r-form field will also be called an r-form for short.

The corresponding vector spaces of r-forms (r-form fields) are called $\Omega^r(M)$:

$$\begin{aligned} &\Omega^0(M) = \mathcal{F}(M) \quad \text{smooth functions on } M \\ &\Omega^1(M) = T^*(M) \quad \text{cotangent vector fields on } M \\ &\Omega^2(M) = sp\{dx^\mu \wedge dx^\nu \mid \mu < \nu\} \\ &\quad \vdots \end{aligned}$$

4.4.11 Exterior derivative

The **exterior derivative** d is a map $\Omega^r(M) \to \Omega^{r+1}(M)$,

$$\omega = \frac{1}{r!}\omega_{\mu_1 \ldots \mu_r} dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r} \mapsto d\omega = \frac{1}{r!}\frac{\partial \omega_{\mu_1 \ldots \mu_r}}{\partial x^\nu} dx^\nu \wedge dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}.$$

Example: dim $M = m = 3$. We have the following r-forms:

- $r = 0: \quad \omega_0 = f(x, y, z)$,
- $r = 1: \quad \omega_1 = \omega_x(x, y, z)dx + \omega_y(x, y, z)dy + \omega_z(x, y, z)dz$,
- $r = 2: \quad \omega_2 = \omega_{xy}(x, y, z)dx \wedge dy + \omega_{yz} dy \wedge dz + \omega_{zx} dz \wedge dx$,
- $r = 3: \quad \omega_3 = \omega_{xyz} dx \wedge dy \wedge dz$.

The exterior derivatives are:

- $d\omega_0 = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$. Thus the components are the components of ∇f.
- $d\omega_1 = \frac{\partial \omega_x}{\partial y} dy \wedge dx + \frac{\partial \omega_x}{\partial z} dz \wedge dx + \frac{\partial \omega_y}{\partial x} dx \wedge dy + \frac{\partial \omega_y}{\partial z} dz \wedge dy + \frac{\partial \omega_z}{\partial x} dx \wedge dz + \frac{\partial \omega_z}{\partial y} dy \wedge dz$
  $= \left(\frac{\partial \omega_y}{\partial x} - \frac{\partial \omega_x}{\partial y}\right) dx \wedge dy + \left(\frac{\partial \omega_z}{\partial y} - \frac{\partial \omega_y}{\partial z}\right) dy \wedge dz + \left(\frac{\partial \omega_x}{\partial z} - \frac{\partial \omega_z}{\partial x}\right) dz \wedge dx$
  These are the components of $\nabla \times \vec{\omega} \quad (\vec{\omega} = (\omega_x, \omega_y, \omega_z))$
- $d\omega_2 = \frac{\partial \omega_{xy}}{\partial dz} dz \wedge dx \wedge dy + \frac{\partial \omega_{yz}}{\partial x} dx \wedge dy \wedge dz + \frac{\partial \omega_{zx}}{\partial y} dy \wedge dz \wedge dx$
  $= \left(\frac{\partial \omega_{yz}}{\partial x} + \frac{\partial \omega_{zx}}{\partial y} + \frac{\partial \omega_{xy}}{\partial z}\right) dx \wedge dy \wedge dz$
  The component is a divergence: $\nabla \cdot \vec{\omega}' \quad$ (where $\vec{\omega}' = (\omega_{yz}, \omega_{zx}, \omega_{xy})$)
- Thus the exterior derivatives correspond to the gradient, curl and divergence! $[d\omega_3 = 0]$

What is $d(d\omega)$?

$$d(d\omega) = \frac{1}{r!}\left(\underbrace{\frac{\partial^2}{\partial x^\alpha \partial x^\beta}}_{\text{symmetric in } \alpha \text{ and } \beta} w_{\mu_1 \ldots \mu_r} \overbrace{dx^\alpha \wedge dx^\beta}^{\text{antisymmetric in } \alpha \text{ and } \beta} \wedge dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r} \right) = 0.$$

So $d^2 = 0$. Note that (for dim $M = 3$)

$$d(df) = d(\partial_x f dx + \partial_y f dy + \partial_z f dz) = \left(\frac{\partial^2}{\partial x \partial y} - \frac{\partial^2}{\partial y \partial x}\right) dx \wedge dy + \ldots = 0,$$

so we recover $\nabla \times \nabla f = 0$. Similarly $d(d\omega_1) = 0 \leftrightarrow \nabla \cdot \nabla \times \vec{\omega} = 0$.

If $d\omega = 0$, we say that ω is a **closed** r-form. If there exists an (r-1)-form ω_{r-1} such that $\omega_r = d\omega_{r-1}$, then we say that ω_r is an **exact** r-form.

The exterior derivative induces the sequence of maps

$$0 \xrightarrow{i} \Omega^0 \xrightarrow{d_0} \Omega^1 \xrightarrow{d_1} \Omega^2 \xrightarrow{d_2} \ldots \xrightarrow{d_{m-2}} \Omega^{m-1} \xrightarrow{d_{m-1}} \Omega^m \xrightarrow{d_m} 0,$$

where $\Omega^r = \Omega^r(M)$, i is the inclusion map $0 \hookrightarrow \Omega^0(M)$ and d_r denotes the map $d_r : \Omega^r \to \Omega^{r+1}$, $\omega \mapsto d\omega$. Since $d^2 = 0$, we have $\underbrace{Im\ d_r}_{\text{exact r+1 forms}} \subset \underbrace{Ker\ d_{r+1}}_{\text{closed r+1 forms}}$. Such a sequence is called an **exact sequence**. This particular sequence is called the **de Rham complex**. The quotient space $Ker\ d_{r+1}/Im\ d_r$ is called the r^{th} **de Rham cohomology group.**

4.4.12 Integration of Differential Forms

Orientable manifolds : Let dim $M = m$. We can define integration over an m-form over M only if M is an **orientable** manifold.
Let $p \in M,\ p \in U_i \cap U_j$ and denote the coordinates on $U_i = \{x^\mu\}$ and on $U_j = \{y^\mu\}$.
T_pM is spanned by $e_\mu = \frac{\partial}{\partial x^\mu}$ or $\tilde{e}_\mu = \frac{\partial}{\partial y^\mu}$. [Recall that $\tilde{e}_\mu = \frac{\partial x^\nu}{\partial y^\mu} e_\nu$ (chain rule)]
Let J denote the determinant $J = \det\left(\frac{\partial x^\mu}{\partial y^\nu}\right)$.
If $J > 0$, we say that $\{e_\mu\}$ and $\{\tilde{e}_\mu\}$ define the **same orientation** on $U_i \cap U_j$.
If $J < 0$, we say that $\{e_\mu\}$ and $\{\tilde{e}_\mu\}$ define the **opposite orientation** on $U_i \cap U_j$.
($J = 0$ is not possible if the coordinates x^μ and y^ν are properly defined.)

We say that $(M, \{U_i, x_i\})$ (manifold M with an atlas $\{U_i, x_i\}$) is **orientable** if for any overlapping charts U_i and U_j the determinant $J = \det\left(\frac{\partial x_i^\mu}{\partial x_j^\nu}\right)$ is positive, $J > 0$. (Note that i and j are fixed, while μ and ν denote the components of the matrix. In other words the determinant is taken over μ and ν.)

If M is orientable, then there exists an m-form ω which is non-vanishing everywhere on M (proof skipped). This m-form ω is called a **volume element** and it plays the role of an integration measure on M. Two volume elements ω and ω' are equivalent, if $\omega = h\omega'$, where $h \in \mathcal{F}$ is a smooth, positive function on M, i.e. $h(p) > 0$ for all $p \in M$. We denote then $\omega \sim \omega'$ (this is clearly an equivalence relation).
If $\omega'' \nsim \omega$, then $\omega = h''\omega''$, where $h''(p) < 0\ \ \forall p \in M$. So there are two equivalence classes for volume elements, corresponding to two inequivalent orientations. We call one of them **right-handed** and the other **left-handed**.

Integration of forms: Let M be orientable and $f : M \to \mathbb{R}$ a function which is nonzero only on one chart $(U_i, x^\mu(p) = \varphi_i^\mu(p))$, and ω a volume element on U_i:

$\omega = h(p)dx^1 \wedge \ldots \wedge dx^m$. We define

$$\int_{U_i} f\omega = \int_{\varphi_i(U_i)} dx^1 dx^2 \ldots dx^m h(\varphi_i^{-1}(x)) f(\varphi_i^{-1}(x))$$

Note that the right hand side is a regular integral in $\mathbb{R}^m$. For a generic function on M, we need to use the "partition of unity".
Let $\{U_i\}$ be an open covering of M, such that every point $p \in M$ belongs to only a finite number of U_i's. (If such an open covering exists, manifold M is called paracompact). The partition of unity is a family of differentiable functions $\epsilon_i(p)$ such that

(i) $0 \leq \epsilon_i(p) \leq 1$

(ii) $\epsilon_i(p) = 0 \ \forall p \notin U_i$

(iii) $\sum_i \epsilon_i(p) = 1 \ \forall p \in M$.

The partition of unity $\{\epsilon_i\}$ depends on the choice of $\{U_i\}$.

Now let $f : M \to \mathbb{R}$. We can write $f(p) = f(p)\sum_i \epsilon_i(p) = \sum_i f_i(p)$, where $f_i = f\epsilon_i$. Then $f_i(p) = 0$ when $p \notin U_i$ so we can use the previous definition to extend the integral over all M:

$$\int_M f\omega = \sum_i \int_{U_i} f_i \omega.$$

Note that due to the paracompactness condition, the sum over i is finite and thus there are no problems with the convergence of the sum. One can show, that although a different atlas $\{(V_i, \psi_i)\}$ gives different coordinates and partition of unity, the integral remains the same.

Example: Let $M = S^1$, $U_1 = S^1 - \{(1,0)\}$, $U_2 = S^1 - \{(-1,0)\}$. Choose the (inverse) coordinate functions as

$$\begin{aligned} \varphi_1^{-1} &: (0, 2\pi) \to U_1, \qquad \theta_1 \mapsto (\cos\theta_1, \sin\theta_1) \\ \varphi_2^{-1} &: (-\pi, \pi) \to U_2, \qquad \theta_2 \mapsto (\cos\theta_2, \sin\theta_2) \end{aligned}$$

Partition of unity: $\epsilon_1(\theta_1) = \sin^2\frac{\theta_1}{2}$, $\epsilon_2(\theta_2) = \cos^2\frac{\theta_2}{2}$. (Note that this satisfies (i) - (iii)). Choose $f : S^1 \to \mathbb{R}$ as $f(\theta) = \sin^2\theta$ and $\omega = 1 \cdot d\theta_1$ on U_1 and $\omega = 1 \cdot d(\theta_2 + 2\pi) = 1 \cdot d\theta_2$ on U_2. Now

$$\int_{S^1} f\omega = \sum_{i=1}^{2} \int_{U_i} f_i\omega = \int_0^{2\pi} d\theta_1 \sin^2\frac{\theta_1}{2} \sin^2\theta_1 + \int_{-\pi}^{\pi} d\theta_2 \cos^2\frac{\theta_2}{2} \sin^2\theta_2 = \frac{\pi}{2} + \frac{\pi}{2} = \pi,$$

as expected.

4.4.13 Lie Groups and Algebras

A **Lie group** G is a differentiable manifold with a group structure,

(i) product $G \times G \to G, \quad (g_1, g_2) \mapsto g_1 g_2$, such that $g_1(g_2 g_3) = (g_1 g_2) g_3$,

(ii) unit element: point $e \in G$ such that $eg = ge = g \ \forall g \in G$,

(iii) inverse element: $\forall g \in G \ \exists g^{-1} \in G$ such that $gg^{-1} = g^{-1} g = e$,

in such a way that the map $G \times G \to G, \quad (g_1, g_2) \mapsto g_1 g_2$ is differentiable. We already know some examples: GL, SL, O, U, SU and SO.

Example: Coordinates on $\mathrm{GL}(n, \mathbb{R}) : x^{ij}(g) = g^{ij}$ (and thus $x^{ij}(e) = \delta^{ij}$.) One chart is sufficient : $U = \mathrm{GL}(n, \mathbb{R})$. (thus U is open in any topology.)

- To be exact we don't yet have a topology on $\mathrm{GL}(n, \mathbb{R})$. We can define the topology in several (inequivalent) ways. One way would be to choose a topology manually, for instance choose the discrete or trivial topology. This is rarely a useful method. A better way of defining the topology is to choose a map f from $\mathrm{GL}(n, \mathbb{R})$ to some known topological space N and then choose the topology on $\mathrm{GL}(n, \mathbb{R})$ so that the map f is continuous, i.e. define

  $$V \subset \mathrm{GL}(n, \mathbb{R}) \text{ is open} \Leftrightarrow V = f^{-1} W \text{ for some } W \text{ open in } N.$$

  (check that this defines a topology). Here are two possible topologies:

  1. Choose $f : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}, \quad g \mapsto \det(g)$. (So we choose $N = \mathbb{R}$). The induced topology is:
     $V \subset \mathrm{GL}(n, \mathbb{R})$ is open $\Leftrightarrow V = f^{-1}(W)$ for some W open in $\mathbb{R}$.
     Note that $\mathrm{GL}(n, \mathbb{R})$ is not Hausdorff with respect to this topology, since if $g_1, g_2 \in \mathrm{GL}(n, \mathbb{R}), \ g_1 \neq g_2$, and $\det g_1 = \det g_2$, then any open set containing g_1 also contains g_2.
  2. Choose $N = \mathbb{R}^{n^2}, \quad f : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^{n^2}$ defined by

     $$\begin{pmatrix} x^{11} & \cdots & x^{1n} \\ \vdots & \ddots & \vdots \\ x^{n1} & \cdots & x^{nn} \end{pmatrix} \mapsto (x^{11}, x^{12}, \ldots, x^{1n}, x^{21}, \ldots, x^{nn}) \in \mathbb{R}^{n^2}.$$

     This is clearly injective, and when we define topology as above, we see that f is a homeomorphism from $\mathrm{GL}(n, \mathbb{R})$ to an open subset of $\mathbb{R}^{n^2}$. Since $\mathbb{R}^{n^2}$ is Hausdorff, so is $\mathrm{GL}(n, \mathbb{R})$ with this topology. Thus this topology is not equivalent to the one defined in the first example. This is the usual topology one has on $\mathrm{GL}(n, \mathbb{R})$.

Let $a \in G$ be a given element. We can define the **left-translation**

$$L_a : G \to G, \quad L_a(g) = ag \quad \text{(group action on itself from the left)}.$$

This is a diffeomorphism $G \to G$.
A vector field X on G is **left-invariant**, if the push satisfies

$$(L_a)_* X|_g = X|_{ag}$$

Using coordinates, this means

$$(L_a)_* X|_g = X^\mu(g) \frac{\partial x^\alpha(ag)}{\partial x^\mu(g)} \frac{\partial}{\partial x^\alpha}\bigg|_{ag} = X|_{ag} = X^\alpha(ag) \left.\frac{\partial}{\partial x^\alpha}\right|_{ag},$$

and thus

$$X^\alpha(ag) = X^\mu(g) \frac{\partial x^\alpha(ag)}{\partial x^\mu(g)}.$$

A left-invariant vector field is uniquely defined by its value at a point, for example at $e \in G$, because

$$X|_g = (L_g)_* X_e \equiv L_{g*} V,$$

where $V = X|_e \in T_e G$. Let's denote the set of left-invariant vector fields by $\mathcal{G}$. It is a vector space (since L_{g*} is a linear map); it is isomorphic with $T_e G$. Thus we have $\dim \mathcal{G} = \dim G$.

Example: The left-invariant fields of $\mathrm{GL}(n, \mathbb{R})$:

$$V = V^{ij} \left.\frac{\partial}{\partial x^{ij}}\right|_e \in T_e \mathrm{GL}(n, \mathbb{R}),$$

$$X|_g = L_{g*} V = V^{ij} \frac{\partial(\overbrace{x^{kl}(g) x^{lm}(e)}^{=x^{km}(g)})}{\partial x^{ij}(e)} \frac{\partial}{\partial x^{km}(g)} = V^{ij} x^{kl}(g) \delta^l_i \delta^m_j \frac{\partial}{\partial x^{km}(g)}$$
$$V^{ij} x^{ki}(g) \frac{\partial}{\partial x^{kj}(g)} = \underbrace{x^{ki}(g) V^{ij}}_{(gV)^{kj}} \frac{\partial}{\partial x^{kj}(g)} = (gV)^{kj} \frac{\partial}{\partial x^{kj}(g)},$$

where V^{ij} is an arbitrary $n \times n$ real matrix.

Since $\mathcal{G}$ is a collection of vector fields, we can compute their commutators. The result is again left-invariant!

$$L_{a*} \left[X, Y\right]\big|_g = [L_{a*} X|_g, L_{a*} Y|_g] \overset{\text{l. inv.}}{=} [X|_{ag}, Y|_{ag}] \equiv [X, Y]|_{ag}.$$

So if $X, Y \in \mathcal{G}$, also $[X, Y] \in \mathcal{G}$.

Definition: The set of left-invariant vector fields $\mathcal{G}$ with the commutator (Lie bracket) $[\ ,\ ] : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ is called the **Lie algebra** of a Lie group G.

Examples:

1. $\mathrm{gl}(n, \mathbb{R}) = n \times n$ real matrices (Lie algebras are written with lower case letters).

2. $\mathrm{sl}(n, \mathbb{R})$: Take a curve $c(t)$ that passes through $e \in \mathrm{SL}(n, \mathbb{R})$ and compute its tangent vector ($c(0) = e = 1_n$). For small t: $c(t) = 1_n + tA$, $\left.\frac{dc}{dt}\right|_{t=0} = A \in T_e\mathrm{SL}(n, \mathbb{R})$. Now $\det c(t) = \det\,(1_n + tA) = 1 + t \,\mathrm{tr}\, A + \ldots = 1$. Thus $\mathrm{tr}\, A = 0$ and $\mathrm{sl}(n, \mathbb{R}) = \{A|\ A$ is a $n \times n$ real matrix, $\mathrm{tr}\ A = 0\}$.

3. $\mathrm{so}(n)$: $c(t) = 1_n + tA$. We need $c(t)$ to be orthogonal:
$c(t)c(t)^T = (1 + tA)(1 + tA^T) = 1 + t(A + A^T) + O(t^2) = 1$. Thus we need to have $A = -A^T$ and so $\mathrm{so}(n) = \{A|\ A$ is an antisymmetric $n \times n$ matrix $\}$.

For complex matrices, the coordinates are taken to be the real and imaginary parts of the matrix

4. $\mathrm{u}(n)$: $c(t) = 1_n + tA$. Thus $c(t)c(t)^\dagger = (1+tA)(1+tA^\dagger) = 1+t(A+A^\dagger)+O(t^2) = 1$. So $A = -A^\dagger$ and $\mathrm{u}(n) = \{A|A$ is an antihermitean $n \times n$ complex matrix $\}$.

<u>Note:</u> In physics, we usually use the convention $c(t) = 1 + itA \Rightarrow A^\dagger = A$
$\Rightarrow \mathrm{u}(n) = \{$Hermitean $n \times n$ matrices $\}$.

5. $\mathrm{su}(n) = \{n \times n$ antihermitean traceless matrices $\}$.

4.4.14 Structure Constants of the Lie Algebra

Let $\{V_1, \ldots, V_n\}$ be a basis of T_eG (assume dim $G = n < \infty$). Then $X_\mu|_g = L_{g*}V_\mu$, $\mu = 1, \ldots, n$ is a basis of T_gG (usually it is not a coordinate basis). Since the vectors $\{V_1, \ldots, V_n\}$ are linearly independent, $\{X_1|_g, \ldots, X_n|_g\}$ are also linearly independent. (L_{g*} is an isomorphism between T_eG and T_gG; $(L_{g*})^{-1} = L_{g^{-1}*}$). Since V_μ are basis vectors of T_eG, we can expand

$$[V_\mu, V_\nu] = c_{\mu\nu}{}^\lambda V_\lambda.$$

Let's then push this to T_gG:

$$\begin{aligned} L_{g*}[V_\mu, V_\nu] &= [L_{g*}V_\mu, L_{g*}V_\nu] = [X_\mu|_g, X_\nu|_g] \\ L_{g*}(c_{\mu\nu}{}^\lambda V_\lambda) &= c_{\mu\nu}{}^\lambda X_\lambda|_g \\ &\Rightarrow [X_\mu|_g, X_\nu|_g] = c_{\mu\nu}{}^\lambda X_\lambda|_g. \end{aligned}$$

Letting g vary over all G, we get the same equation everywhere on G with the same numbers $c_{\mu\nu}{}^\lambda$. Thus we can write

$$[X_\mu, X_\nu] = c_{\mu\nu}{}^\lambda X_\lambda.$$

The $c_{\mu\nu}{}^{\lambda}$ are called the **structure constants** of the Lie algebra. Evidently we have $c_{\mu\nu}{}^{\lambda} = -c_{\nu\mu}{}^{\lambda}$. We also have the Jacobi identity (of commutators)

$$c_{\mu\nu}{}^{\tau} c_{\tau\rho}{}^{\sigma} + c_{\nu\rho}{}^{\tau} c_{\tau\mu}{}^{\sigma} + c_{\rho\mu}{}^{\tau} c_{\tau\nu}{}^{\sigma} = 0.$$

4.4.15 The adjoint representation of G

Let b be some element of G, $b \in G$. Let us define the map

$$ad_b : G \to G, \quad ad_b(g) \equiv ad_b g = bgb^{-1}.$$

This is a homomorphism: $ad_b g_1 \cdot ad_b g_2 = ad_b(g_1 g_2)$, and at the same time defines an action of G on itself (conjugation): $ad_b \cdot ad_c = ad_{bc}, \;\; ad_e = id_G$. (Note that this is really a combined map: $ad_b \cdot ad_c \equiv ad_b \circ ad_c$). The differential map ad_{b*} pushes vectors from $T_g G$ to $T_{ad_b g} G$. If $g = e$, $ad_b e = beb^{-1} = e$, so ad_{b*} maps $T_e G$ to itself. Lets denote this map by Ad_b:

$$Ad_b : T_e G \to T_e G, \quad Ad_b = ad_{b*}|_{T_e G}$$

One can easily show that $(f \circ g)_* = f_* \circ g_*$, thus $ad_{b*} ad_{c*} = ad_{bc*}$. It then follows that Ad_b is a representation of G in the vector space $\mathcal{G} \cong T_e G$, the so-called **adjoint representation**:

$$Ad : G \to \mathrm{Aut}(\mathcal{G}), \quad b \mapsto Ad_b.$$

If G is a matrix group (O, SO,...), then $V \in T_e G \cong \mathcal{G}$ is a matrix and

$$Ad_g V = gVg^{-1}.$$

(This follows from $ad_g(e + tV) = e + tgVg^{-1}$.) So, if $\{V_\mu\}$ is a basis of $\mathcal{G}$,

$$gV_\mu g^{-1} = V_\nu D^{(\mathrm{adj})\nu}{}_{\mu}(g).$$

4.5 Integral of an r-form over a manifold M; Stokes' theorem

4.5.1 Simplexes in a Euclidean space

We define simplexes in $\mathbb{R}^m$ as follows:

0-simplex : point $s^0 = p_0$
1-simplex : oriented line $s^1 = (p_0, p_1)$
2-simplex : oriented triangle $s^2 = (p_0, p_1, p_2)$
3-simplex : oriented tetrahedron $s^3 = (p_0, p_1, p_2, p_3)$
$\vdots$

n-simplex $(p_0, \ldots, p_n)$ is made of (n+1) geometrically independent[2] points (vertices) $p_0, \ldots, p_n$ in this order and the n-dimensional object spanned by them:

$$s^n = \{x \in \mathbb{R}^m | x^\mu = \sum_{i=0}^{n} t_i x^\mu(p_i),\ \sum_{i=0}^{n} t_i = 1,\ t_i \geq 0\}$$

The numbers $t_0, \ldots, t_n$ are the barycentric coordinates on s^n.
As a subset of $\mathbb{R}^m$ s^n is closed and bounded and therefore compact. The orientation is defined by the order of the vertices. If $\Pi \in S_{n+1}$ is a permutation of (n+1)-elements, then we define

$$(p_{\Pi(0)}, \ldots, p_{\Pi(n)}) = (-1)^\Pi (p_0, \ldots, p_n),$$

so even permutations of the vertices give the same oriented simplex s^n, and odd permutations give the simplex $-s^n$ with opposite orientation.
The **boundary** ∂s^n of an n-simplex s^n is a combination of (n-1)-simplexes: If $s^n = (p_0, \ldots, p_n)$,

$$\partial s^n = \sum_{i=0}^{n} (-1)^i (p_0, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n).$$

Example: $\partial s^0 = 0$

$$\begin{aligned}
&s^1 = (p_0, p_1), && \partial s^1 = p_1 - p_0 \\
&s^2 = (p_0, p_1, p_2), && \partial s^2 = (p_1, p_2) - (p_0, p_2) + (p_0, p_1) = (p_1, p_2) + (p_0, p_1) + (p_2, p_0) \\
&s^3 = (p_0, p_1, p_2, p_3), && \partial s^3 = (p_1, p_2, p_3) - (p_0, p_2, p_3) + (p_0, p_1, p_3) - (p_0, p_1, p_2) \\
& && \quad = (p_1, p_2, p_3) + (p_0, p_3, p_2) + (p_0, p_1, p_3) + (p_1, p_0, p_2).
\end{aligned}$$

An **n-chain** c is a formal sum

$$c = \sum_i a^i s_i^n, \qquad a^i \in \mathbb{R},\ s_i^n \text{ an n-simplex.}$$

Thus ∂s^n is an (n-1)-chain. The boundary of the chain is: $\partial c \equiv \sum_i a^i \partial s_i^n$. A boundary has no boundary, so we should have $\partial^2 c = 0$. Let us prove this. It is enough to prove this for a simplex since ∂ is defined as a linear operator.

$$\partial^2 s^n = \partial \left(\sum_{i=0}^{n} (-1)^i (p_0, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n) \right)$$

Let $j < k$. In $\partial^2 s^n$ the simplex $(p_0, \ldots, p_{j-1}, p_{j+1}, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n)$ is created in two ways:

1. The first ∂ removes p_k and the second p_j: sign $(-1)^{k+j}$
2. The first ∂ removes p_j and the second p_k: sign $(-1)^{j+(k-1)}$.

[2]Geometrically independent $\equiv$ vectors $p_0 - p_1, \ldots, p_0 - p_n$ are linearly independent and thus span an n-dimensional space.

Thus the two terms have opposite signs and cancel each other $\Rightarrow \partial^2 s^n = 0$.

Two n-simplexes, $P = (p_0, \ldots, p_n)$ and $Q = (q_0, \ldots, q_n)$, can be mapped onto each other with an orientation preserving linear homeomorphism. The image of $p \in P$ in Q is the point with the same barycentric coordinates t_i.

In $\mathbb{R}^m$ we define the **standard simplex** $\bar{s}^m = (p_0, \ldots, p_m)$ as follows:

$$\begin{aligned} p_0 &= (0, 0, \ldots, 0) \quad \text{(origin)} \\ p_1 &= (1, 0, \ldots, 0) \\ p_2 &= (0, 1, \ldots, 0) \\ &\vdots \\ p_m &= (0, 0, \ldots, 1). \end{aligned}$$

Now let ω be an m-form on $U \subset \mathbb{R}^m$, where $\bar{s}^m \subset U$. Now ω can be written as

$$\omega = A(x^1, x^2, \ldots, x^m) dx^1 \wedge dx^2 \wedge \ldots \wedge dx^m.$$

Let us define the integral of ω over the standard simplex:

$$\int_{\bar{s}^m} \omega \equiv \int_{\bar{s}^m} dx^1 \ldots dx^m A(x^1, \ldots, x^m).$$

Example: Consider $m = 3, \quad \omega = dx \wedge dy \wedge dz$:

$$\begin{aligned} \int_{\bar{s}^3} \omega &= \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz = \int_0^1 dx \int_0^{1-x} dy (1 - x - y) \\ &= \int_0^1 dx ((1-x)^2 - \frac{1}{2}(1-x)^2) = \frac{1}{2} \int_0^1 dx (1-x)^2 = \frac{1}{6} \end{aligned}$$

4.5.2 Simplexes and Chains on Manifolds

Let M be a manifold of dimension m and $s^n \subset U \subset \mathbb{R}^n$ a Euclidean n-simplex ($s^n = (p_0, \ldots, p_n)$). In addition $\varphi : U \to M$ is a smooth map (does not need to be injective or surjective) where U is open. A "protosimplex" on M is (s^n, U, φ). If $t^n = (q_0, \ldots, q_n) \subset V \subset \mathbb{R}^m$ is another Euclidean n-simplex and $\psi : V \to M$, then $(s^n, U, \varphi) \sim (t^n, V, \psi)$ if

$$\psi(\sum_{i=0}^{n} t^i x^\mu(q_i)) = \varphi(\sum_{i=0}^{n} t^i x^\mu(p_i))$$

with the same t^i. (So the points with the same barycentric coordinates map to the same point on M). We can see that $\sim$ is an equivalence relation.

An n-simplex σ^n on M is an equivalence class in the equivalence relation above. If (s^n, U, φ) is a representative of σ^n and the "sides" of s^n are $t_0, \ldots, t_n$: $\partial s^n = \sum \pm t_i$, then the sides of σ^n are $\tau_i = (t_i, V_i, \varphi)$, where $t_i \subset V_i \subset U$ (V_i open in $\mathbb{R}^{n-1}$) and the

boundary of σ^n is $\partial\sigma^n = \sum \pm\tau_i$.
An n-chain on M is a formal sum $c = \sum a_i\sigma_i^n$, where $a_i \in \mathbb{R}$ and σ_i^n is an n-simplex. Addition of chains is defined by $\alpha c + \beta c' \equiv \sum_i(\alpha a_i + \beta a_i')\sigma_i^n$. The boundary of the chain is $\partial c \equiv \sum a_i \partial\sigma_i^n$.

If we denote by $C_n(M)$ the set of chains ($C_n(M) = \{$ n-chains on $M\}$), then we have a linear map $\partial : C_n(M) \to C_{n-1}(M)$ with the property $\partial^2 = 0$. A **cycle** z is a chain with a vanishing boundary: $\partial z = 0$. (Compare with closed n-forms : $d\omega = 0$). A cycle b is a **boundary cycle** or **boundary** if there exists an (n+1)-chain c such that $b = \partial c$. (Compare with exact n-forms: $\omega = d\alpha$ for some (n-1)-form α). Every boundary is a cycle, but not vice versa. (Compare with all exact forms are closed but not vice versa).

Integration of Forms Let M be a manifold, ω a p-form on M and c a p-chain on M. We wish to define

$$\int_c \omega.$$

Let us write $c = \sum_i a_i s_i$, where s_i's are p-simplexes, and let us define

$$\int_c \omega = \sum_i a_i \int_{s_i} \omega.$$

This means that we have to define the integral of ω over a simplex s. We can write the simplex in the form $(\bar{s}^p, U, \varphi)$, where $\bar{s}^p$ is a standard simplex in $\mathbb{R}^p$, $\varphi : U \to M$, $\bar{s}^p \subset U$. Now we can define

$$\int_s \omega \equiv \int_{\bar{s}^p} \varphi^*\omega.$$

In practice there are often more practical methods to calculate.

Stokes' Theorem: Let $\omega \in \Omega^{r-1}(M)$ and c be an r-chain on M. Then

$$\int_c d\omega = \int_{\partial c} \omega.$$

Proof: Due to linearity it is enough to show this for a simplex: $\int_s d\omega = \int_{\partial s} \omega$. Writing s as $(\bar{s}^r, U, \varphi)$ we can write

$$\int_s d\omega = \int_{\bar{s}^r} \varphi^*(d\omega) \stackrel{*}{=} \int_{\bar{s}^r} d(\varphi^*\omega),$$

where $(*)$ is an exercise. Similarly

$$\int_{\partial s} \omega = \int_{\partial \bar{s}^r} \varphi^*\omega.$$

Thus it is enough to show that in $\mathbb{R}^r$ we have

$$\int_{\bar{s}^r} d\eta = \int_{\partial \bar{s}^r} \eta, \qquad \eta \in \Omega^{r-1}(\mathbb{R}^r).$$

In general $\eta = \sum_\mu a_\mu(x) dx^1 \wedge \ldots \wedge dx^{\mu-1} \wedge dx^{\mu+1} \wedge \ldots \wedge dx^r$. It is enough to examine one term, for instance $\eta = a(x) dx^1 \wedge \ldots \wedge dx^{r-1}$. Then $d\eta = (-1)^{r-1} \frac{\partial a(x)}{\partial x^r} dx^1 \wedge \ldots \wedge dx^r$. A direct calculation gives

$$\begin{aligned}\int_{\bar{s}^r} d\eta &= (-1)^{r-1} \int_{\bar{s}^r} \frac{\partial a(x)}{\partial x^r} dx^1 \ldots dx^r \\ &= (-1)^{r-1} \int_{x^\mu \geq 0, \sum x^\mu \leq 1} dx^1 \ldots dx^{r-1} \int_0^{1-\sum_{\mu=1}^{r-1}} dx^r \frac{\partial a(x)}{\partial x^r} \\ &= (-1)^{r-1} \int dx^1 \ldots dx^{r-1} \left(a(x^1, \ldots, x^{r-1}, 1 - \sum_{\mu=1}^{r-1} x^\mu) - a(x^1, \ldots, x^{r-1}, 0) \right)\end{aligned} \tag{12}$$

Now $\partial \bar{s}^r = (p_1, \ldots, p_r) - (p_0, p_2, \ldots, p_r) + \ldots + (-1)^r (p_0, \ldots, p_{r-1})$. The sides $(p_0, p_2, \ldots, p_r), \ldots, (p_0, p_1, \ldots, p_{r-2}, p_r)$ are all subsets of the planes $x^\mu = 0, \quad \mu = 1, 2, \ldots, r-1$. In the plane $x^\mu = 0$ the μ component of vectors is zero, i.e. $\eta(v_1, \ldots, v_{r-1}) = 0$. Therefore on these sides $\eta = 0$, only sides $(p_1, \ldots, p_r)$ and $(-1)^r (p_0, \ldots, p_{r-1})$ contribute. The latter part is a standard simplex:

$$(-1)^r \int_{(p_0, \ldots, p_{r-1})} \eta = (-1)^r \int_{\bar{s}^{r-1}} dx^1 \ldots dx^{r-1} a(x^1, \ldots, x^{r-1}, 0).$$

This is the second term in (12). $\sigma \equiv (p_1, \ldots, p_r)$ is not a standard simplex. The integral over it is defined by mapping σ to a standard simplex preserving orientation. This is done by mapping points with the same barycentric coordinates to each other, which here simply means a projection to the $x^r = 0$ plane:

$$(p_1, \ldots, p_{r-1}, p_r) \mapsto (p_1, \ldots, p_{r-1}, p_0) = (-1)^{r-1} (p_0, \ldots, p_{r-1}) = (-1)^{r-1} \bar{s}^{r-1}.$$

Therefore

$$\int_{(p_1, \ldots, p_r)} \eta = (-1)^{r-1} \int_{\bar{s}^{r-1}} dx^1 \ldots dx^{r-1} a(x^1, \ldots, x^{r-1}, 1 - \sum x^\mu)$$

This is the first term in (12). Therefore $\int_c d\omega = \int_{\partial c} \omega$. □

5 Riemannian Geometry (Metric Manifolds)

(Chapter 7 of Nakahara's book)

5.1 The Metric Tensor

Let M be a differentiable manifold. The Riemannian metric on M is a $(0,2)$-tensorfield, which satisfies

(i) $g_p(U,V) = g_p(V,U) \qquad \forall p \in M, \;\; U,V \in T_pM$ (i.e. g is *symmetric*)

(ii) $g_p(U,U) \geq 0$, and $g_p(U,U) = 0 \Leftrightarrow U = 0$ (g is *positive definite*).

If instead of (ii) g satisfies

(ii') If $g_p(U,V) = 0$ for all $U \in T_pM$, then $V = 0$,

we say that g is a pseudo-Riemannian metric (symmetric and non-degenerate).
(M,g) with a (pseudo-) Riemannian metric is called a (pseudo-) Riemannian manifold. The spacetime in general relativity is an example of a pseudo-Rimannian manifold. In local coordinates $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$. (The Euclidean metric: $g_{\mu\nu} = \delta_{\mu\nu}$. Then $g(U,V) = \sum_{i=1}^n U^i V^i$.)

5.2 The Induced Metric

Let (N, g_N) be a Riemannian manifold, dim $N = n$. We define an m dimensional **submanifold** M of N:
Let $f : M \to N$ be a smooth map such that f is an injection and the push $f_* : T_pM \to T_{f(p)}N$ is also an injection. Then f is an **embedding** of M in N and the image $f(M)$ is a **submanifold** of N. However, it follows that M and $f(M)$ are diffeomorphic, so we can call M a submanifold of N.

Now the pullback f^* of f induces the natural metric g_M on M:

$$g_M = f^* g_N.$$

The components of g_M are given by

$$g_{M\mu\nu}(x) = g_{N\alpha\beta}(f(x)) \frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\beta}{\partial x^\nu}.$$

[By the chain rule: $g_{M\mu\nu} dx^\mu \otimes dx^\nu = g_{N\alpha\beta} \frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\beta}{\partial x^\nu} dx^\mu \otimes dx^\nu$]

Example: Let (θ, φ) be the polar coordinates on S^2 and $f : S^2 \to \mathbb{R}^3$ the usual embedding: $f(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. On $\mathbb{R}^3$ we have the Euclidean metric $\delta_{\mu\nu}$. We denote $y^1 = \theta, y^2 = \varphi$. We obtain the induced metric on S^2:

$$g_{\mu\nu} dy^\mu \otimes dy^\nu = \delta_{\alpha\beta} \frac{\partial f^\alpha}{\partial y^\mu} \frac{\partial f^\beta}{\partial y^\nu} dy^\mu \otimes dy^\nu = d\theta \otimes d\theta + \sin^2\theta d\varphi \otimes d\varphi.$$

Thus the components of the metric are $g_{11}(\theta,\varphi) = 1, \;\; g_{22}(\theta,\varphi) = \sin^2\theta, \;\; g_{12}(\theta,\varphi) = g_{21}(\theta,\varphi) = 0$.

Why the notation ds^2 is often used for the metric?

Often the metric is denoted $ds^2 = g_{\mu\nu} dx^\mu \otimes dx^\nu$. The reason for this is as follows. Let $c(t)$ be a curve on manifold M with the metric g. The tangent vector of the curve is $\dot{c}(t)$, which in local coordinates is $\dot{c}(t) = (\frac{dx^\mu(t)}{dt})$. $[c(t) = (x^\mu(t))]$

If $M = \mathbb{R}^3$ with the Euclidean metric $g_{\mu\nu} = \delta_{\mu\nu}$, the length of the curve between t_0 and t_1 would be

$$L_{\mathbb{R}^3} = \int_{t_0}^{t_1} dt \sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2 + (\dot{x}^3)^2} = \int_{t_0}^{t_1} dt \sqrt{\delta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}.$$

In general case the length of the part of the curve between t_0 and t_1 is then

$$L = \int_{t_0}^{t_1} dt \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}. \tag{13}$$

If t_0 and t_1 are infinitesimally close : $t_1 = t_0 + \Delta t$, then

$$\Delta s \equiv L \approx \Delta t \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} \approx \Delta t \sqrt{g_{\mu\nu} \frac{\Delta x^\mu}{\Delta t} \frac{\Delta x^\nu}{\Delta t}} = \sqrt{g_{\mu\nu} \Delta x^\mu \Delta x^\nu}.$$

Thus $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ is the square of an "infinitesimal length element" ds. We will have more to say about (13) later.

5.3 Affine Connection

Recall that $\chi(M) = \{$ vector fields on $M\}$. An (affine) **connection** ∇ is a map $\chi(M) \times \chi(M) \to \chi(M)$, $(X, Y) \mapsto \nabla_X Y$ such that

1. $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$ (linear in the 2^{nd} argument)
2. $\nabla_{(X+Y)} Z = \nabla_X Z + \nabla_Y Z$ (linear in the 1^{st} argument)
3. f is a function on M $(f \in \mathcal{F}(M)) \Rightarrow \nabla_{fX} Y = f \nabla_X Y$
4. $\nabla_X (fY) = X[f]Y + f \nabla_X Y$.

Now take a chart (U, φ) with coordinates $x = \varphi(p)$. Let $\{e_\nu = \frac{\partial}{\partial x^\nu}\}$ be the coordinate basis of $T_p M$. We define $(\dim M)^3$ **connection coefficients** $\Gamma^\lambda{}_{\mu\nu}$ by

$$\nabla_{e_\mu} e_\nu = \Gamma^\lambda{}_{\mu\nu} e_\lambda.$$

We can express the connection in the coordinate basis with the help of connection coefficients: Let $X = X^\mu e_\mu$ and $Y = Y^\nu e_\nu$ be two vector fields. Denote $\nabla_\mu \equiv \nabla_{e_\mu}$. Now

$$\begin{aligned} \nabla_X Y &\overset{2,3}{=} X^\mu \nabla_\mu (Y^\nu e_\nu) \overset{4}{=} X^\mu e_\mu [Y^\nu] e_\nu + X^\mu Y^\nu \nabla_\mu e_\nu = X^\mu \frac{\partial Y^\nu}{\partial x^\mu} e_\nu + X^\mu Y^\nu \Gamma^\lambda{}_{\mu\nu} e_\lambda \\ &= X^\mu (\frac{\partial Y^\lambda}{\partial x^\mu} + \Gamma^\lambda{}_{\mu\nu} Y^\nu) e_\lambda \equiv X^\mu (\nabla_\mu Y)^\lambda e_\lambda, \end{aligned}$$

where we have

$$(\nabla_\mu Y)^\lambda = \frac{\partial Y^\lambda}{\partial x^\mu} + \Gamma^\lambda{}_{\mu\nu} Y^\nu.$$

Note that $\nabla_X Y$ contains no derivatives of X unlike $\mathcal{L}_X Y$.

5.4 Parallel Transport and Geodesics

Let $c : (a, b) \to M$ be a curve on M with coordinate representation $x^\mu = x^\mu(t)$. Its tangent vector is

$$V = V^\mu e_\mu|_{c(t)} = \frac{dx^\mu(c(t))}{dt} e_\mu\bigg|_{c(t)}.$$

If a vector field X satisfies

$$\nabla_V X = 0 \qquad \text{(along } c(t)\text{)},$$

then we say that X is **parallel transported** along the curve $c(t)$. In component form this is

$$\frac{dX^\mu}{dt} + \Gamma^\mu{}_{\nu\lambda} \frac{dx^\nu(t)}{dt} X^\lambda = 0.$$

If the tangent vector V itself is parallel transported along the curve $c(t)$,

$$\nabla_V V = 0, \tag{14}$$

then the curve $c(t)$ is called a **geodesic**. The equation (14) is the **geodesic equation** and in component form it is

$$\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu{}_{\nu\lambda} \frac{dx^\nu}{dt} \frac{dx^\lambda}{dt} = 0$$

Geodesics can be interpreted as the straightest possible curves in a Riemannian manifold. If $M = \mathbb{R}^n$ and $\Gamma = 0$, then the geodesics are straight lines.

5.5 The Covariant Derivative of Tensor Fields

Connection was a term that we used for the map $\nabla : (X, Y) \mapsto \nabla_X Y$. The map $\nabla_X : \chi(M) \to \chi(M)$, $Y \mapsto \nabla_X Y$ is called the covariant derivative. It is a proper generalization of the directional derivative of functions to vector fields, and as we'll discuss next, to tensor fields.

For a function, we define $\nabla_X f$ to be the same as the directional derivative:

$$\nabla_X f = X[f].$$

Thus the condition number 4 in the definition of ∇ is the Leibnitz rule:

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

Let's require that this should be true for any product of tensors:

$$\nabla_X(T_1 \otimes T_2) = (\nabla_X T_1) \otimes T_2 + T_1 \otimes (\nabla_X T_2),$$

where T_1 and T_2 are tensor fields of arbitrary types. The formula must also be true when some of the indices are contracted. Thus we can define the covariant derivative of a one-form as follows. Let $\omega \in \Omega^1(M)$ be a one-form ((0,1) tensor field), $Y \in \chi(M)$ be a vector field ((1,0) tensor field). Then $< \omega, Y > \in \mathcal{F}(M)$ is a smooth function on M. Recall that $< \omega, Y > \equiv \omega[Y] = \omega_\mu Y^\mu$. (Here μ is the contracted index.) Then

$$\nabla_X < \omega, Y > = X(\omega[Y]) = X^\mu \frac{\partial}{\partial x^\mu}(\omega_\nu Y^\nu) = X^\mu \frac{\partial \omega_\nu}{\partial x^\mu} Y^\nu + X^\mu \omega_\nu \frac{\partial Y^\nu}{\partial x^\mu}.$$

On the other hand because of the Leibnitz rule we must have

$$\begin{aligned} \nabla_X < \omega, Y > &= < \nabla_X \omega, Y > + < \omega, \nabla_X Y > = (\nabla_X \omega)_\nu Y^\nu + \omega_\nu (\nabla_X Y)^\nu \\ &= (\nabla_X \omega)_\nu Y^\nu + \omega_\nu X^\mu \frac{\partial Y^\nu}{\partial x^\mu} + \omega_\nu \Gamma^\nu{}_{\mu\alpha} X^\mu Y^\alpha \end{aligned}$$

From these two formulas we find $(\nabla_X \omega)_\nu$. (Note that the two $X^\mu \omega_\nu \frac{\partial Y^\nu}{\partial x^\mu}$ terms cancel.)

$$\Rightarrow (\nabla_X \omega)_\nu = X^\mu \left(\frac{\partial \omega_\nu}{\partial x^\mu} - \Gamma^\alpha{}_{\mu\nu} \omega_\alpha \right).$$

When $X = \frac{\partial}{\partial x^\mu}$, this reduces to

$$(\nabla_\mu \omega)_\nu = \frac{\partial \omega_\nu}{\partial x^\mu} - \Gamma^\alpha{}_{\mu\nu} \omega_\alpha.$$

Further when $\omega = dx^\sigma$: $\nabla_\mu dx^\sigma = -\Gamma^\sigma{}_{\mu\nu} dx^\nu$.
For a generic tensor, the result turns out to be

$$\begin{aligned} (\nabla_\nu t)^{\lambda_1 \ldots \lambda_p}_{\mu_1 \ldots \mu_q} &= \partial_\nu t^{\lambda_1 \ldots \lambda_p}_{\mu_1 \ldots \mu_q} + \Gamma^{\lambda_1}{}_{\nu\rho} t^{\rho\lambda_2 \ldots \lambda_p}_{\mu_1 \ldots \mu_q} + \ldots + \Gamma^{\lambda_p}{}_{\nu\rho} t^{\lambda_1 \ldots \lambda_{p-1}\rho}_{\mu_1 \ldots \mu_q} \\ &- \Gamma^\rho{}_{\nu\mu_1} t^{\lambda_1 \ldots \lambda_p}_{\rho\mu_2 \ldots \mu_q} - \ldots - \Gamma^\rho{}_{\nu\mu_q} t^{\lambda_1 \ldots \lambda_p}_{\mu_1 \ldots \mu_{q-1}\rho}. \end{aligned}$$

(Note that we should really have written $t^{\lambda_1 \ldots \lambda_p}{}_{\mu_1 \ldots \mu_q}$, but this was not done for typographical reasons.)

5.6 The Transformation Properties of Connection Coefficients

Let U and V be two overlapping charts with coordinates:

$$\begin{aligned} &\text{on } U: \quad x \quad e_\mu = \frac{\partial}{\partial x^\mu}, \\ &\text{on } V: \quad y \quad \tilde{e}_\nu = \frac{\partial}{\partial y^\nu} = \frac{\partial x^\mu}{\partial y^\nu} e_\mu. \end{aligned}$$

Let $p \in U \cap V \neq \emptyset$. The connection coefficients on V are

$$\nabla_{\tilde{e}_\alpha} \tilde{e}_\beta = \tilde{\Gamma}^\gamma{}_{\alpha\beta} \tilde{e}_\gamma = \tilde{\Gamma}^\gamma{}_{\alpha\beta} \frac{\partial x^\nu}{\partial y^\gamma} e_\nu$$

On the other hand

$$\nabla_{\tilde{e}_\alpha} \tilde{e}_\beta = \nabla_{\tilde{e}_\alpha} (\frac{\partial x^\mu}{\partial y^\beta} e_\mu) = \left(\frac{\partial^2 x^\nu}{\partial y^\alpha y^\beta} + \frac{\partial x^\lambda}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\beta} \Gamma^\nu{}_{\lambda\mu} \right) e_\nu$$

Thus

$$\tilde{\Gamma}^\gamma{}_{\alpha\beta} \frac{\partial x^\nu}{\partial y^\gamma} = \left(\frac{\partial^2 x^\nu}{\partial y^\alpha y^\beta} + \frac{\partial x^\lambda}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\beta} \Gamma^\nu{}_{\lambda\mu} \right).$$

From this we find the transformation rule for the connection coefficients:

$$\tilde{\Gamma}^\gamma{}_{\alpha\beta} = \frac{\partial y^\gamma}{\partial x^\nu} \frac{\partial x^\lambda}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\beta} \Gamma^\nu{}_{\lambda\mu} + \frac{\partial^2 x^\nu}{\partial y^\alpha \partial y^\beta} \frac{\partial y^\gamma}{\partial x^\nu}.$$

We notice that the first term is just the transformation rule for the components of a (1,2)-tensor. But we also have an additional second term, which is symmetric in α and β. Thus Γ is <u>almost</u> like a (1,2)-tensor, but not quite. To construct a (1,2)-tensor out of Γ, define

$$T^\gamma{}_{\alpha\beta} = \Gamma^\gamma{}_{\alpha\beta} - \Gamma^\gamma{}_{\beta\alpha} \equiv 2\Gamma^\gamma{}_{[\alpha\beta]} = \text{ the } \textbf{torsion tensor}$$

(note: $t_{[\alpha\beta]} = \frac{1}{2}(t_{\alpha\beta} - t_{\beta\alpha})$ is the antisymmetrization of indices.)

5.7 The Metric Connection

Let c be an arbitrary curve and V its tangent vector. If a connection ∇ satisfies[3]

$$\nabla_V(g(X,Y)) = 0 \qquad \text{when} \quad \nabla_V X = 0 \text{ and } \nabla_V Y = 0,$$

then we say that ∇ is a **metric connection.** Since

$$\nabla_V(g(X,Y)) = (\nabla_V g)(X,Y) + g(\overbrace{\nabla_V X}^{=0}, Y) + g(X, \overbrace{\nabla_V Y}^{=0}) = 0,$$

the metric connection satisfies

$$\nabla_V g = 0.$$

In component form:

1. $(\nabla_\mu g)_{\alpha\beta} = \partial_\mu g_{\alpha\beta} - \Gamma^\lambda{}_{\mu\alpha} g_{\lambda\beta} - \Gamma^\lambda{}_{\mu\beta} g_{\alpha\lambda} = 0.$
   And by cyclic permutation of μ, α and β we get:

[3]This condition means that the angle between vectors is preserved under parallel transport.

2. $(\nabla_\alpha g)_{\beta\mu} = \partial_\alpha g_{\beta\mu} - \Gamma^\lambda{}_{\alpha\beta} g_{\lambda\mu} - \Gamma^\lambda{}_{\alpha\mu} g_{\beta\lambda} = 0$

3. $(\nabla_\beta g)_{\mu\alpha} = \partial_\beta g_{\mu\alpha} - \Gamma^\lambda{}_{\beta\mu} g_{\lambda\alpha} - \Gamma^\lambda{}_{\beta\alpha} g_{\mu\lambda} = 0$

Let us denote the symmetrization of indices: $\Gamma^\gamma{}_{(\alpha\beta)} \equiv \frac{1}{2}(\Gamma^\gamma{}_{\alpha\beta} + \Gamma^\gamma{}_{\beta\alpha})$. Then adding -(1)+(2)+(3) gives

$$-\partial_\mu g_{\alpha\beta} + \partial_\alpha g_{\beta\mu} + \partial_\beta g_{\mu\alpha} + T^\lambda{}_{\mu\alpha} g_{\lambda\beta} + T^\lambda{}_{\mu\beta} g_{\lambda\alpha} - 2\Gamma^\lambda{}_{(\alpha\beta)} g_{\lambda\mu} = 0$$

In other words

$$\Gamma^\lambda{}_{(\alpha\beta)} g_{\lambda\mu} = \frac{1}{2}\left\{(\partial_\alpha g_{\beta\mu} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta}) + T^\lambda{}_{\mu\alpha} g_{\lambda\beta} + T^\lambda{}_{\mu\beta} g_{\lambda\alpha}\right\}$$

Thus

$$\Gamma^\kappa{}_{(\alpha\beta)} = \left\{ \begin{matrix} \kappa \\ \alpha\beta \end{matrix} \right\} + \frac{1}{2}(T_\alpha{}^\kappa{}_\beta + T_\beta{}^\kappa{}_\alpha),$$

where $\left\{ \begin{smallmatrix} \kappa \\ \alpha\beta \end{smallmatrix} \right\} = \frac{1}{2} g^{\kappa\mu}(\partial_\alpha g_{\beta\mu} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta})$ are the **Christoffel symbols** and $T_\alpha{}^\kappa{}_\beta = g_{\alpha\lambda} g^{\kappa\mu} T^\lambda{}_{\mu\beta}$.

The coefficients of a metric connection thus satisfy

$$\Gamma^\kappa{}_{\alpha\beta} = \Gamma^\kappa{}_{(\alpha\beta)} + \Gamma^\kappa{}_{[\alpha\beta]} = \left\{ \begin{matrix} \kappa \\ \alpha\beta \end{matrix} \right\} + \underbrace{\frac{1}{2}\left(T_\alpha{}^\kappa{}_\beta + T_\beta{}^\kappa{}_\alpha + T^\kappa{}_{\alpha\beta}\right)}_{\equiv K^\kappa{}_{\alpha\beta} = \textbf{ contorsion}}.$$

If the torsion tensor vanishes, $T^\kappa{}_{\alpha\beta} = 0$, the metric connection is called the **Levi-Civita connection**:

$$\Gamma^\kappa{}_{\alpha\beta} = \left\{ \begin{matrix} \kappa \\ \alpha\beta \end{matrix} \right\}.$$

5.8 Curvature And Torsion

We define two new tensors:

(Riemann) curvature tensor: $R : \chi(M) \times \chi(M) \times \chi(M) \to \chi(M)$

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

Torsion tensor: $T : \chi(M) \times \chi(M) \to \chi(M)$

$T(X, Y) \equiv \nabla_X Y - \nabla_Y X - [X, Y].$

Let's check that these definitions really define tensors, i.e. multilinear maps. Obviously $R(X + X', Y, Z) = R(X, Y, Z) + R(X', Y, Z)$ etc. are true, but it is less obvious that $R(fX, gY, hZ) = fghR(X, Y, Z)$ where $f, g, h \in \mathcal{F}(M)$. Let's calculate:

$$[fX, gY] = fX[g]Y - gY[f]X + fg[X, Y] \tag{15}$$

Using (15) we obtain

$$\begin{aligned} R(fX, gY)(hZ) = {} & f\nabla_X(g\nabla_Y(hZ)) - g\nabla_Y(f\nabla_X(hZ)) \\ & - fX[g]\nabla_Y(hZ) + gY[f]\nabla_X(hZ) - fg\nabla_{[X,Y]}(hZ). \end{aligned}$$

Here the first term is

$$f\nabla_X(g\nabla_Y(hZ)) = f\nabla_X(gY[h]Z + gh\nabla_Y Z) = fX[g]Y[h]Z + fg(X[Y[h]])Z \\ + fgY[h]\nabla_X Z + fgX[h]\nabla_Y Z + fhX[g]\nabla_Y Z + fgh\nabla_X\nabla_Y Z,$$

and the second term is obtained by changing $X \leftrightarrow Y$ and $f \leftrightarrow g$. Continuing

$$\begin{aligned} R(fX, gY)(hZ) = & fX[g]Y[h]Z + fg(X[Y[h]])Z + fgY[h]\nabla_X Z + fgX[h]\nabla_Y Z \\ & + fhX[g]\nabla_Y Z + fgh\nabla_X\nabla_Y Z - gY[f]X[h]Z - fg(Y[X[h]])Z \\ & - fgX[h]\nabla_Y Z - fgY[h]\nabla_X Z - ghY[f]\nabla_X Z - fgh\nabla_Y\nabla_X Z \\ & - fX[g]Y[h]Z - fhX[g]\nabla_Y Z + gY[f]X[h]Z + ghY[f]\nabla_X Z \\ & - fg([X,Y][h])Z - fgh\nabla_{[X,Y]}Z = fgh(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z) \\ & = fghR(X,Y)Z. \end{aligned}$$

Thus R is a linear map. In other words, when $X = X^\mu e_\mu, Y = Y^\nu e_\nu$ and $Z = Z^\lambda e_\lambda$, we have

$$R(X,Y)Z = X^\mu Y^\nu Z^\lambda R(e_\mu, e_\nu)e_\lambda.$$

R maps three vector fields to a vector field, so it is a (1,3)-tensor. A similar (but shorter) calculation shows that $T(fX, gY) = fgT(X,Y)$, so $T(X,Y) = X^\mu Y^\nu T(e_\mu, e_\nu)$. T is a (1,2) tensor.

The operations of R and T on vectors are obtained by knowing their actions on the basis vectors $e_\mu \frac{\partial}{\partial x^\mu}$. Denote

$$R(e_\mu, e_\nu)e_\lambda = \text{ a vector, expand in basis } e_\kappa = R^\kappa{}_{\lambda\mu\nu}e_\kappa.$$

Note the placement of indices. We can derive a formula for obtaining the components $R^\kappa{}_{\lambda\mu\nu}$. Recall that $[e_\mu, e_\nu] = 0$ and $dx^\kappa(e_\sigma) = \delta^\kappa{}_\sigma$. Thus we get

$$\begin{aligned} R^\kappa{}_{\lambda\mu\nu} &= dx^\kappa(R(e_\mu, e_\nu)e_\lambda) = dx^\kappa(\nabla_\mu\nabla_\nu e_\lambda - \nabla_\nu\nabla_\mu e_\lambda) = dx^\kappa(\nabla_\mu(\Gamma^\eta{}_{\nu\lambda}e_\eta) - \nabla_\nu(\Gamma^\eta{}_{\mu\lambda}e_\eta)) \\ &= dx^\kappa((\partial_\mu\Gamma^\eta{}_{\nu\lambda})e_\eta) + \Gamma^\eta{}_{\nu\lambda}\Gamma^\rho{}_{\mu\eta}e_\rho - (\partial_\nu\Gamma^\eta{}_{\mu\lambda})e_\eta - \Gamma^\eta{}_{\mu\lambda}\Gamma^\rho{}_{\nu\eta}e_\rho) \end{aligned} \tag{16}$$

Therefore

$$\boxed{R^\kappa{}_{\lambda\mu\nu} = \partial_\mu\Gamma^\kappa{}_{\nu\lambda} - \partial_\nu\Gamma^\kappa{}_{\mu\lambda} + \Gamma^\eta{}_{\nu\lambda}\Gamma^\kappa{}_{\mu\eta} - \Gamma^\eta{}_{\mu\lambda}\Gamma^\kappa{}_{\nu\eta}}$$

Similarly if we denote $T(e_\mu, e_\nu) = T^\lambda{}_{\mu\nu}e_\lambda$ and derive the components $T^\lambda{}_{\mu\nu}$:

$$T^\lambda{}_{\mu\nu} = dx^\lambda(T(e_\mu, e_\nu)) = dx^\lambda(\nabla_\mu e_\nu - \nabla_\nu e_\mu) = dx^\lambda(\Gamma^\eta{}_{\mu\nu}e_\eta - \Gamma^\eta{}_{\nu\mu}e_\eta),$$

and therefore

$$T^\lambda{}_{\mu\nu} = \Gamma^\lambda{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\mu}.$$

Thus this is the <u>same</u> torsion tensor as the one we had defined earlier.

Geometric interpretation:

Let us also define:
The Ricci tensor: $Ric(X,Y) = dx^\lambda(R(e_\lambda, Y)X)$. Thus the components are:
$(Ric)_{\mu\nu} = Ric(e_\mu, e_\nu) = R^\lambda{}_{\mu\lambda\nu}$. (Usual notation $(Ric)_{\mu\nu} \equiv R_{\mu\nu}$.)
The scalar curvature: $R = g^{\mu\nu}(Ric)_{\mu\nu} = R^{\lambda\nu}{}_{\lambda\nu}$.
The Einstein tensor: $G_{\mu\nu} = (Ric)_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$.

5.9 Geodesics of Levi-Civita Connections

The length of a curve $c(s) = (x^\mu(s))$ is defined by

$$I(c) = \int_c ds = \int_c \sqrt{g_{\mu\nu}\frac{dx^\mu}{ds'}\frac{dx^\nu}{ds'}}ds' \equiv \int_c L ds'$$

Thus along a curve L is constant. One can normalize s' such that $L = 1$ so $s' = s$. Curves with extremal (minimum or maximum) length satisfy $\delta I = 0$ about the curve. (Variational principle.) They satisfy the Euler-Lagrange equations (familiar from calculus of variations (FYMM II)):

$$\frac{d}{ds}\left(\frac{\partial L}{\partial x'^\mu}\right) - \frac{\partial L}{\partial x^\mu} = 0, \qquad \text{where } x'^\mu = \frac{dx^\mu}{ds} \tag{17}$$

L = Lagrange function or Lagrangian. Instead of L, which contains a square root, we can equivalently use a simpler Lagrange function

$$F = \frac{1}{2}g_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = \frac{1}{2}L^2,$$

because

$$\frac{d}{ds}\left(\frac{\partial F}{\partial x'^\mu}\right) - \frac{\partial F}{\partial x^\mu} = L\underbrace{\left(\frac{d}{ds}\left(\frac{\partial L}{\partial x'^\mu}\right) - \frac{\partial L}{\partial x^\mu}\right)}_{=0} + \frac{\partial L}{\partial x'^\mu}\underbrace{\frac{dL}{ds}}_{=0} = 0,$$

when $x^\mu(s)$ satisfies the Euler-Lagrange equation (17). Then $\delta(\int F ds) = 0$ gives

$$\begin{aligned}
&\frac{d}{ds}\left(g_{\lambda\mu}\frac{dx^\mu}{ds}\right) - \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^\lambda}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 0 \\
&\Rightarrow \frac{\partial g_{\lambda\mu}}{\partial x^\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} + g_{\lambda\mu}\frac{d^2x^\mu}{ds^2} - \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^\lambda}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 0 \\
&\Rightarrow g_{\lambda\mu}\frac{d^2x^\mu}{ds^2} + \frac{1}{2}\left(\frac{\partial g_{\lambda\mu}}{\partial x^\nu} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\lambda}\right)\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 0.
\end{aligned}$$

Multiply this by $g^{\kappa\lambda}$ and sum over λ:

$$\boxed{\frac{d^2x^\kappa}{ds^2} + \left\{\begin{matrix}\kappa \\ \mu\nu\end{matrix}\right\}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 0.} \tag{18}$$

This is the geodesic equation with a Levi-Civita connection! The action $I = \int F ds$ sometimes provides a convenient starting point for computing the Christoffel symbols $\left\{ {\kappa \atop \mu\nu} \right\}$: plug in the metric to I, derive the Euler- Lagrange equations and read off the Christoffel symbols comparing the Euler-Lagrange equations with (18).
Note: previously when we discussed the geodesic equation in the context of general connection, we said that geodesics are the "straightest" possible curves. Now, in the context of the Levi-Civita connection which is only based on the metric, we that the geodesics are also the shortest possible curves.

Note also that we can explicitly restore a parameter m and write the action of the length of the curve as $I = m \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{ds'} \frac{dx^\nu}{ds'}} ds'$. This is the relativistic action of a free massive point particle (with mass m) moving on a curved spacetime. Thus the free point particles move along geodesics. If $m^2 > 0$ (usual particles), we say that the corresponding geodesics (on a pseudo-Riemannian manifold) are **timelike**, if $m^2 < 0$ (tachyonic particles) the geodesics are **spacelike**. Massless particles (such as the photon) move along **null** geodesics. The invariant length vanishes along a null geodesic, $ds^2 = 0$. This equation can be used to determine the null geodesics.

5.10 Lie Derivative And the Covariant Derivative

Let $\Gamma^\mu{}_{\nu\lambda}$ be an arbitrary symmetric ($\Gamma^\mu{}_{\nu\lambda} = \Gamma^\mu{}_{\lambda\nu}$) connection. We can then re-express the Lie derivative with the help of the covariant derivative as follows:

$$(\mathcal{L}_X Y)^\mu = X^\nu \partial_\nu Y^\mu - Y^\nu \partial_\nu X^\mu = X^\nu \nabla_\nu Y^\mu - (\nabla_\nu X^\mu) Y^\nu$$

This is true because of the symmetry of the connection:

$$\begin{aligned} X^\nu \nabla_\nu Y^\mu - (\nabla_\nu X^\mu) Y^\nu &= X^\nu (\partial_\nu Y^\mu + \Gamma^\mu{}_{\nu\lambda} Y^\lambda) - (\partial_\nu X^\mu + \Gamma^\mu{}_{\nu\lambda} X^\lambda) Y^\nu \\ & X^\nu \partial_\nu Y^\mu - Y^\nu \partial_\nu X^\mu + (\underbrace{\Gamma^\mu{}_{\nu\lambda} - \Gamma^\mu{}_{\lambda\nu}}_{=0}) X^\nu Y^\lambda \end{aligned}$$

For a generic (p,q)-tensor:

$$\begin{aligned} \mathcal{L}_X T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q} &= (X^\lambda \nabla_\lambda) T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q} - (\nabla_\lambda X^{\mu_1}) T^{\lambda \mu_2 \ldots \mu_p}_{\nu_1 \ldots \nu_q} - \ldots - (\nabla_\lambda X^{\mu_p}) T^{\mu_1 \ldots \mu_{p-1} \lambda}_{\nu_1 \ldots \nu_q} \\ &+ (\nabla_{\nu_1} X^\lambda) T^{\mu_1 \ldots \mu_p}_{\lambda \nu_2 \ldots \nu_q} + \ldots + (\nabla_{\nu_q} X^\lambda) T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_{q-1} \lambda}. \end{aligned}$$

5.11 Isometries

Isometries are a very important concept. They are symmetries of a Riemannian manifold. If the manifold is a spacetime, we usually require a physical theory to be invariant under isometries.

Definition. Let (M, g) be a (pseudo)-Riemannian manifold. A diffemorphism $f : M \to M$ is an **isometry** if it preserves the metric,

$$f^* g_{f(p)} = g_p \ ,$$

for all $p \in M$.

If we interpret the metric as a map on vector fields, the above requirement means

$$g_{f(p)}(f_* X, f_* Y) = g_p(X, Y) \tag{19}$$

for all tangent vectors $X, Y \in T_p M$. In component form, (19) is

$$\frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} \ g_{\alpha\beta}(f(p)) = g_{\mu\nu}(p) \tag{20}$$

where x, y are coordinates of the points $p, f(p)$ respectively. What (19) means, is that an isometry must preserve the angles between all tangent vectors and their lengths.

The identity map is trivially an isometry, also the composite map $f \circ g$ of two isometries f, g is an isometry. Further, if f is an isometry, so is its inverse f^{-1}. This means that isometries form a group with composition of maps as the product, called the **isometry group**. The isometry group is a group of symmetries of a (pseudo)-Riemannian manifold.

Examples.

- $(M, g) =$ the Euclidean space (R^n, δ) with the Euclidean metric. All translations $x^\mu \mapsto x^\mu + a^\mu$ in some direction $a = (a^\mu)$ are isometries, and so are rotations. The isometry group $\{$translations, rotations, and their combinations$\}$ is called the **Euclidean group** or **Galilean group** and denoted by E^n.

- $(M, g) =$ the $(d+1)$-dimensional Minkowski space(time) $(R^{1,d}, \eta)$ with the Minkowski metric η. Again, spacetime translations $x^\mu \mapsto x^\mu + a^\mu$ are isometries, additional isometries are (combinations of these and) space rotations and boosts. The isometry group $\{$translations, rotations, boosts, and their combinations$\}$ is called the **Poincaré** group.

In typical laboratory scales, our spacetime is approximately flat (a Minkowski space) so its approximate isometry group is the Poincaré group. That's the reason for special relativity and the requirement that physics in the laboratory be relativistic, *i.e.* Poincaré invariant. More precisely, that requirement is necessary for experiments which involve scales where relativistic effects become important. For lower scales, time "decouples" and we can make a further approximation where only the Euclidean isometries of the spacelike directions are relevant. Recall also that symmetries such as the time translations and space translations lead into conservation laws, like the conservation of energy and momentum. As you can see, important physical principles are a reflection of the isometries of the spacetime.

5.12 Killing Vector Fields

Let us now consider the limit of "small" isometries, *i.e.* infinitesimal displacements $x = p \mapsto f(p) = y \approx x + \epsilon X$. Here ϵ is an infinitesimal parameter and X is a vector field indicating the direction of the infinitesimal displacement. If the above map is an isometry, the vector field X is called a **Killing vector field**. Since the infitesimal displacement is an isometry, eqn. (20) must be satisfied and it now takes the form

$$\frac{\partial(x^\alpha + \epsilon X^\alpha)}{\partial x^\mu} \frac{\partial(x^\beta + \epsilon X^\beta)}{\partial x^\nu} \; g_{\alpha\beta}(x + \epsilon X) = g_{\mu\nu}(x) \tag{21}$$

By Taylor expanding the left hand side, and requiring that the leading infinitesimal term of order ϵ vanishes (there's no ϵ-dependence on the right hand side), we obtain the equation

$$X^\xi \partial_\xi g_{\mu\nu} + \partial_\mu X^\alpha g_{\alpha\nu} + \partial_\nu X^\beta g_{\mu\beta} = 0 \; . \tag{22}$$

We can recognize the left hand side as a Lie derivative, so (22) can be rewritten as

$$\mathcal{L}_X g_{\mu\nu} = 0 \; .$$

Expressing $\mathcal{L}_X g_{\mu\nu}$ with the help of the covariant derivative,

$$\mathcal{L}_X g_{\mu\nu} = X^\lambda \overbrace{\nabla_\lambda g_{\mu\nu}}^{=0} + (\nabla_\mu X^\lambda) g_{\lambda\nu} + (\nabla_\nu X^\lambda) g_{\mu\lambda} = 0.$$

($\nabla_\lambda g_{\mu\nu} = 0$) for a metric connection). Thus a Killing vector field satisfies

$$\boxed{\nabla_\mu X_\nu + \nabla_\nu X_\mu = 0} \qquad \text{Killing equation.}$$

Let X and Y be two Killing vector fields. We can easily verify that

a) all linear combinations $aX + bY$ with $a, b \in R$ are also Killing vector fields

b) the Lie bracket $[X, Y]$ is a Killing vector field

It then follows that the Killing vector fields form an algebra, the **Lie algebra of the isometry group**. (The isometry group is usually a Lie group.)

Now let $x^\mu(t)$ be a geodesic, its tangent vector $U^\mu = \frac{dx^\mu}{dt}$, and let V^μ be a Killing vector. Then,

$$(U^\nu \nabla_\nu)(U^\mu V_\mu) = \underbrace{U^\mu U^\nu \nabla_\nu V_\mu}_{=\frac{1}{2} U^\mu U^\nu (\nabla_\mu V_\nu + \nabla_\nu V_\mu)} + V_\mu \underbrace{U^\nu \nabla_\nu U^\mu}_{=0 \;\; \text{(geodesic)}} = 0.$$

Thus $U^\mu V_\mu = U \cdot V$ is a *constant on a geodesic.*

An m-dimensional manifold M can have at most $\frac{1}{2}m(m+1)$ linearly independent Killing vector fields. Manifolds with the maximum number of Killing vector fields are called **maximally symmetric**. E.g. $\mathbb{R}^m$ is maximally symmetric ($g_{\mu\nu} = \delta_{\mu\nu} \Rightarrow \Gamma = 0$). The Killing equation $\partial_\mu V_\nu + \partial_\nu V_\mu = 0$ has solutions:

$$\begin{aligned} &V^\mu_{(i)} = \delta^\mu_i \quad (m \text{ of these}) \\ &V_\mu = a_{\mu\nu} x^\nu \quad \text{with} \quad \underbrace{a_{\mu\nu} = -a_{\nu\mu}}_{\frac{1}{2}m(m-1) \text{ components}} \quad = \text{ constant } \neq 0 \end{aligned} \tag{23}$$

Thus in total we have $m + \frac{1}{2}m(m-1) = \frac{1}{2}m(m+1)$. Ok.