# BACKGROUND STUDY FOR PARTICLE PHYSICS

Sanjay Rijal

**Start Date: February 15, 2023**

# 1 LINEAR ALGEBRA

## 1.1 Background Definitions

A vector space **V** over a field $\mathbb{F}$ is a set {u, v, w, ...} of vectors, together with a set {a, b, c, ...} of elements in $\mathbb{F}$ called scalars, that is closed under the taking of linear combinations:

$$u,\ v\ \epsilon\ \mathbf{V}\ and\ a,\ b\ \epsilon\ \mathbb{F} \to au\ +\ bv\ \epsilon\ \mathbf{V}, \tag{1}$$

and where 0v = 0 and 1v = v

A subspace of **V** is a subset of **V** that is also a **vector space**. An **affine subspace** of **V** is a translate of a subspace of **V**. The vector space **V** is the direct sum of two subspaces **U** and **W**, written $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$, if $\mathbf{U} \cap \mathbf{W} = 0$ (the only vector in common is the zero vector) and every vector v $\epsilon$ **V** can be written uniquely as v = u + w for some u $\epsilon$ **U**and w $\epsilon$ **W**.

A set $\{v_i\}$ of vectors is **linearly independent** (over the field $\mathbb{F}$) if, for any collection of scalars $\{c_i\} \subset \mathbb{F}$,

$$\sum_i c_i v_i = 0 \text{ implies } c_i = 0\ \forall\ i. \tag{2}$$

Essentially this means that no member of a set of linearly independent vectors may be expressed as a linear combination of the others.

A set **B** of vectors is a **spanning set** for **V** (or, more simply, spans **V** ) if every vector in **V** can be written as a linear combination of vectors from **B**. A spanning set of linearly independent vectors is called a basis for the vector space. The cardinality of a basis for **V** is called the dimension of the space, written dim **V**.

Let **V** and **W** be vector spaces. A map $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ is linear (or a homeomorphism) if, $\forall\ v_1, v_2\ \epsilon\ \mathbf{V}$and $a_1, a_2\ \epsilon\ \mathbb{F}$,

$$\mathrm{T}(a_1 v_1 + a_2 v_2) = a_1 \mathrm{T}\ v_1 + a_2 \mathrm{T}\ v_2 \tag{3}$$

The set of all v $\epsilon$ **V** such that T v = 0 is called the **kernel** (or null space) of T, written ker T; dim ker T is sometimes called the nullity of T. The set of all w $\epsilon$ **W** for which there exists a v $\epsilon$ **V** with T v = w is called the image (or range) of T, written im T. The rank of T, rk T, is defined as dim im T.

If T is bijective it is called an **isomorphism**, in which case **V** and **W** are said to be isomorphic; this is written as $\mathbf{V} \cong \mathbf{W}$ or, sloppily, **V** = **W** . Isomorphic vector spaces are not necessarily identical, but they behave as if they were. A linear map from a vector space to itself is called an **endomorphism**, and if it is a bijection it is called an **automorphism**. (Physicists tend to call an endomorphism a linear operator) A linear map T is idempotent if $\mathrm{T}^2 = \mathrm{T}$ . An idempotent endomorphism $\pi$: $\mathbf{V} \to \mathbf{V}$ is called a projection (operator). Remark: This is not to be confused with an orthogonal projection, which requires an inner product for its definition.

**Group homeomorphisms (Linear Mapping)**

- **Monomorphism:** Injective Mapping (one-to-one)
- **Epimorphism:** Surjective Mapping (onto)
- **Isomorphism:** Bijective Mapping (one-to-one and onto)
- **Endomorphism:** Linear Mapping from vector to itself (same domain and co-domain)
- **Automorphism:** Bijective Endomorphism (isomorphic endomorphism)
- **Automorphism:** Isomorphism of Smooth Manifolds (bijection where both function and its inverse are differential)

## 1.2 Dual Space

A linear functional on $\mathbf{V}$ is a linear map f : $\mathbf{V} \to \mathbb{F}$. The set $\mathbf{V}^*$ of all linear functionals on $\mathbf{V}$ is called the **dual space** of $\mathbf{V}$, and is often denoted as Hom($\mathbf{V}$, $\mathbb{R}$). If f is a linear functional and a is a scalar, a f is another linear functional, defined by (a f)(v) = a f(v) (pointwise multiplication). Also, if f and g are two linear functionals then we can obtain a third linear functional f + g by (f + g)(v) = f(v) + g(v) (pointwise addition). These two operations turn $\mathbf{V}^*$ into a vector space, and when one speaks of the dual space one always has this vector space structure in mind. It is customary to write <v, f> or <f, v> to denote f (v). When written this way it is called the **natural pairing** or dual pairing between $\mathbf{V}$ and $\mathbf{V}^*$ . Elements of $\mathbf{V}^*$ are often called covectors.

A **Covector** is simply a linear function from vectors to real numbers, $\alpha$:$\mathbf{V} \to \mathbb{R}$. For an example of a covector, we have these functions $dx_i$ which is not a length but a function that takes vectors and picks out the $i^{th}$ coordinate component, for example in $\mathbb{R}^3$:

$$\begin{aligned} dx_1(A\hat{e_1} + B\hat{e_2} + C\hat{e_3}) &= A \\ dx_2(A\hat{e_1} + B\hat{e_2} + C\hat{e_3}) &= B \\ dx_3(A\hat{e_1} + B\hat{e_2} + C\hat{e_3}) &= C \end{aligned} \tag{4}$$

These functions form a basis in the space of covectors. Every linear function from vectors to the real numbers can be written as a linear combination of these functions:

$$\alpha(v) = \sum_i \alpha_i dx_i(v) \tag{5}$$

If {e$_i$} is a basis of $\mathbf{V}$ , there is a canonical dual basis or **cobasis** $\{\theta_j\}$ of $\mathbf{V}**$ , defined by e$_i$ , $\theta_j = \delta_i j$ , where $\delta_i j$ is the **Kronecker delta**:

$$\delta_i j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$$

Any element f $\epsilon$ $\mathbf{V}^*$ can be expanded in terms of the dual basis as:

$$f = \sum_i f_i \theta_i \tag{6}$$

where f$_i$ $\epsilon$ $\mathbb{F}$. The scalars f$_i$ are called the components of f with respect to the basis $\{\theta_i\}$.

## 1.3 Inner Product

Let $\mathbb{F}$ be a subfield of $\mathbb{C}$, and let $\mathbf{V}$ be a vector space over $\mathbb{F}$. A **sesquilinear form** on $\mathbf{V}$ is a map $g : \mathbf{V} \times \mathbf{V} \to \mathbb{F}$ satisfying the following two properties. For all u, v, w $\epsilon$ $\mathbf{V}$ and a, b $\epsilon$ $\mathbb{F}$, the map $g$ is

1. **linear on the second entry:** $g(\mathrm{u}, \mathrm{av} + \mathrm{bw}) = \mathrm{a}g(\mathrm{u}, \mathrm{v}) + \mathrm{b}g(\mathrm{u}, \mathrm{w})$, and
2. **Hermitian:** $g(\mathrm{v}, \mathrm{u}) = \overline{\mathrm{g}(\mathrm{u}, \mathrm{v})}$ where $\bar{\mathrm{a}}$ is the complex conjugate of a.

These two properties together imply that $g$ is **antilinear** on the first entry: $g(\mathrm{au} + \mathrm{bv}, \mathrm{w}) = \bar{\mathrm{a}}g(\mathrm{u}, \mathrm{w}) + \bar{\mathrm{b}}g(\mathrm{v}, \mathrm{w})$.

However, if $\mathbb{F}$ is a real field (a subfield of $\mathbb{R}$) then $\bar{\mathrm{a}} = \mathrm{a}$ and $\bar{\mathrm{b}} = \mathrm{b}$ and the above condition condition just says that $g$ is linear on the first entry as well. In that case we say that $g$ is a **bilinear form**. Moreover, the Hermiticity condition becomes the symmetry condition $g(\mathrm{u}, \mathrm{v}) = g(\mathrm{v}, \mathrm{u})$, so a real sesquilinear form is in fact a **symmetric bilinear form**. If the sesquilinear form g is

3. **nondegenerate**, so that $g(\mathrm{u}, \mathrm{v}) = 0$ for all v implies $\mathrm{u} = 0$,

then it is called an **inner product**. A vector space equipped with an inner product is called an **inner product space**.

# 2 MULTILINEAR ALGEBRA

## 2.1 Tensors

Multilinear algebra is just the linear algebra with many vector spaces ath the same time. The fundamental objects are **tensors** instead of vectors. A tensor is any **multilinear map** from a vector space to a scalar field. Tensors are not generalizations of vectors in any way. It's very slightly more understandable to say that tensors are generalizations of matrices, in the same way that it is slightly more accurate to say "vanilla ice cream is a generalization of chocolate ice cream" than it is to say that "vanilla ice cream is a generalization of dessert", closer, but still false. Vanilla and Chocolate are both ice cream, but chocolate ice cream is not a type of vanilla ice cream, and "dessert" certainly isn't a type of vanilla ice cream. This definition as a multilinear maps is another reason people think tensors are generalization of matrices, because matrices are linear maps just like tensors. But the distinction is that matrices take a vector space to itself, while tensors take a vector space to a scalar field. So a matrix is not strictly speaking a tensor.

Essentially, one can view a tensor either passively as an element of a certain vector space (the tensor product space) or actively as a multilinear functional on the dual of that vector space. The map

$$\mathbf{T} = \underbrace{\mathbf{V} \times \mathbf{V} \times ... \times \mathbf{V}}_{r\ times} \times \underbrace{\mathbf{V}^* \times \mathbf{V}^* \times ... \times \mathbf{V}^*}_{s\ times} \to \mathbb{R} \tag{7}$$

is said to be multilinear if it is linear in each entry:

$$\mathbf{T}(v_1, ..., au + bw, ..., v_{r+s}) = a\mathbf{T}(v_1, ..., u, ..., v_{r+s}) + b\mathbf{T}(v_1, ..., w, ..., v_{r+s}) \tag{8}$$

The space of all such maps is linear under pointwise addition and scalar multiplication.

Given two vectors v and w, we can form their **tensor product** v $\otimes$ w. The product v $\otimes$ w is called a tensor of order 2 or a second-order tensor or a 2-tensor. It is a vector space to which is associated a bilinear map $\mathbf{V} \times \mathbf{W} \to \mathbf{V} \otimes \mathbf{W}$ that maps a pair (v, w), v $\epsilon$ $\mathbf{V}$ and w $\epsilon$ $\mathbf{W}$, to an element of $\mathbf{V} \otimes \mathbf{W}$, denoted by v$\otimes$w.

**Properties:**

- If $\mathbf{R}$ is a tensor of order $r$ and $\mathbf{S}$ is a tensor of order $s$, then order of $\mathbf{R}\otimes\mathbf{S} = r+s$.
- **Scalar Associativity:** $\mathbf{T} \otimes (a\mathbf{S}) = (a\mathbf{T}) \otimes \mathbf{S} = a\,(\mathbf{T}\otimes\mathbf{S})$ ; a = scalar
- **Associativity:** $(\mathbf{R}\otimes\mathbf{S}) \otimes \mathbf{T} = \mathbf{R}\otimes (\mathbf{S}\otimes\mathbf{T})$
- **Distributive:** $\mathbf{R}\otimes (\mathbf{S} + \mathbf{T}) = \mathbf{R}\otimes\mathbf{S} + \mathbf{R}\otimes\mathbf{T}$
- **Distributive:** $(\mathbf{R} + \mathbf{S}) \otimes (\mathbf{T} + \mathbf{U}) = (\mathbf{R}\otimes\mathbf{T}) + (\mathbf{R}\otimes\mathbf{U}) + (\mathbf{S}\otimes\mathbf{T}) + (\mathbf{S}\otimes\mathbf{U})$
- **Non-Commutative:** $\mathbf{R}\otimes\mathbf{S} \neq \mathbf{S}\otimes\mathbf{R}$
- dim $(\mathbf{R}\otimes\mathbf{S})$ = dim $\mathbf{R}$ . dim $\mathbf{S}$ = dim Hom(R, S)

Let $\mathbf{V}$ be a vector space and $\mathbf{V}^*$ be its dual space. Then a Tensor $\mathbf{T}$ of **type*(r, s)*** is an element of the tensor product space

$$\mathbf{T}^r_s = \underbrace{\mathbf{V} \otimes \mathbf{V} \otimes ... \otimes \mathbf{V}}_{r\ times} \otimes \underbrace{\mathbf{V}^* \otimes \mathbf{V}^* \otimes ... \otimes \mathbf{V}^*}_{s\ times} = \mathbf{V}^{\otimes r} \otimes (\mathbf{V}^*)^{\otimes s} \tag{9}$$

What we previously called a tensor of order $r$ is just a tensor of type *(r, 0)*. The properties of the tensor product ensure that the space of all tensors forms a **multigraded algebra**.

## 2.2 Symmetry Types of Tensors

Let $\mathbf{T}$ be a tensor of type (0, 2) with components $\mathbf{T}_{ij}$ in some basis. If $\mathbf{T}_{ij} = \mathbf{T}_{ji}$ we say $\mathbf{T}$ is **symmetric**, while if $\mathbf{T}_{ij}$ = -$\mathbf{T}$ ji we say $\mathbf{T}$ is **antisymmetric**. Of course, a general (0, 2) tensor has no such property. But if a (0, 2) tensor has either property, we say it has **definite symmetry**.

The symmetric part of T is the symmetric tensor with components,

$$(T_{sym})_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) \tag{10}$$

while the antisymmetric part of T is the antisymmetric tensor with components,

$$(T_{asym})_{ij} = \frac{1}{2}(T_{ij} - T_{ji}) \tag{11}$$

Evidently, for a (0, 2) tensor, $\mathbf{T}$ is the sum of its symmetric and antisymmetric parts. These ideas can be generalized to higher-order tensors, but the results are not as simple. Defining what is meant by "definite symmetry" for higher-order tensors requires some understanding of the representation theory of the symmetric group. The higher-order analogues of symmetric and antisymmetric tensors can be given by,

$$(T_{sym})_{i_1,i_2,...,i_p} = \frac{1}{p!}\sum_{\sigma \epsilon \mathbf{S}_p} T_{i_\sigma(1),i_\sigma(2),...,i_\sigma(p)} = T_{(i_1 i_2 ... i_p)} \tag{12}$$

$$(T_{asym})_{i_1,i_2,...,i_p} = \frac{1}{p!}\sum_{\sigma \epsilon \mathbf{S}_p} (-1)^\sigma T_{i_\sigma(1),i_\sigma(2),...,i_\sigma(p)} = T_{[i_1 i_2 ... i_p]} \tag{13}$$

where $\mathbf{S}_p$ is the set of all permutations of p elements and (-1)$^\sigma$ denotes the sign of the permutation $\sigma$. If p>2 then it is no longer true that T is the sum of its symmetric and antisymmetric parts. symmetry type is preserved under the taking of linear combinations of tensors of the same symmetry type; thus, the set of all symmetric and the set of all antisymmetric tensors are both subspaces of the space of all tensors. Specifically, we write Sym$^p$ $\mathbf{V}$ and Alt$^p$ $\mathbf{V}$ for the subspace of symmetric and antisymmetric respectively (0, p) tensors or ( p, 0) tensors.

Viewed as a multilinear map, the condition that T be symmetric is just

$$T(v_{\sigma(1)}, v_{\sigma(2)}, ..., v_{\sigma(p)}) = T(v_1, v_2, ..., v_p), \tag{14}$$

while the condition that T be antisymmetric is

$$T(v_{\sigma(1)}, v_{\sigma(2)}, ..., v_{\sigma(p)}) = (-1)^{\sigma} T(v_1, v_2, ..., v_p), \tag{15}$$

for any collection of p vectors ($v_1$, $v_2$, ..., $v_p$) and any permutation $\sigma\epsilon\, \mathbf{S}_p$.

## 2.3 Exterior Algebra

Exterior algebra or Grassmann algebra uses exterior/wedge product as multiplication. The exterior product of two vectors u and v also known as **bivector** is given by,

$$u \bigwedge v = u \otimes v - v \otimes u \tag{16}$$

The wedge product turns the collection $\bigwedge \mathbf{V}$ of all $\bigwedge^p$ $\mathbf{V}$ for p = 0, 1, 2, . . . into a graded algebra, called the exterior algebra of $\mathbf{V}$. The wedge product of two vectors is also called **2-vector or 2-blade** and the vector space generated by the set of 2-vectors is denoted as $\bigwedge^2 \mathbf{V}$. For any vector u, v,w $\epsilon \mathbf{V}$,

1. $v \bigwedge v = 0$ (like cross-product)
2. $v \bigwedge w$ = - $(w \bigwedge v)$ (like cross-product)
3. $(u \bigwedge v) \bigwedge w = u \bigwedge (v \bigwedge w)$ (unlike cross-product)

The 2-blade can be interpreted as the area pf the parallelogram with sides of the length equal to the magnitude of vectors, which in 3D can be interpreted as the cross-product of two vectors. In general, all parallel plane surfaces with the same orientation and area have the same bivector as a measure of their oriented area. The exterior product of k-vectors (**k-blade**) lies in the space, $k^{th}$ exterior power and the magnitude of k-blade in general gives the oriented hypervolume of the $k^{th}$ dimensional parallelotope whose edges are the given vectors.

### 2.3.1 Alternating Tensors and the Space $\bigwedge^p$f p-vectors

$\bigwedge^2$ $\mathbf{V}$ is naturally isomorphic to the vector space $\mathrm{Alt}^2\mathbf{V}$ of alternating (2, 0) tensors. First note that $\bigwedge^2$ $\mathbf{V}$ is spanned by all 2-vectors of the form $e_i \bigwedge$ $e_j$ where i < j.

Let us choose a basis $e_1$ , $e_2$ , . . . , $e_n$ for $\mathbf{V}$ . If

$$v = \sum_i v^i e_i \qquad \text{and} \qquad w = \sum_j w^j e_j \tag{17}$$

then by linearity of tensor product,

$$v \bigwedge w = \sum_{ij} (v^i e_i \otimes w^j e_j - w^j e_j \otimes v^i e_i) = \sum_{ij} v^i w^j e_i \bigwedge e_j \tag{18}$$

$$\sum_{ij} v^i w^j e_i \bigwedge e_j = \sum_{ij} v^j w^i e_j \bigwedge e_i = -\sum_{ij} v^j w^i e_i \bigwedge e_j \tag{19}$$

Therefore we can write (again using the antisymmetric property of the wedge product)

$$v \bigwedge w = \frac{1}{2} \sum_{ij} (v^i w^j - w^j v^i) e_i \otimes e_j = \sum_{i<j} (v^i w^j - w^j v^i) e_i \otimes e_j \tag{20}$$

It follows that any linear combination of 2-vectors can be written as a linear combination of the elements $e_i \bigwedge e_j$ for i < j. Moreover, these 2-vectors are all linearly independent, so they form a basis for the space of 2-vectors. There are ${}^nC_2$ such vectors with,

$$dim(\bigwedge^2)\mathbf{V} =^n C_2 = dim(Alt^2\mathbf{V}) \tag{21}$$

where n = dim($\mathbf{V}$)

In general, for k-blade,

$$dim(\bigwedge^k)\mathbf{V} =^n C_k = dim(Alt^k\mathbf{V}) \tag{22}$$

and

$$v_1 \bigwedge v_2 \bigwedge ... \bigwedge v_p = c_p \sum_{\sigma \epsilon \mathbf{S}_p} (-1)^\sigma \mathbf{V}_{\sigma(1)} \otimes \mathbf{V}_{\sigma(2)} \otimes ... \otimes \mathbf{V}_{\sigma(p)} \tag{23}$$

where $c_p$ is some constant (the choice is mostly irrelevant). It is much easier to deal with wedge products than with alternating sums such as those on the right-hand side of equation 23.

### 2.3.2 Hodge Dual

**Hodge star operator** or **Hodge star** is a linear map defined on the exterior algebra of finite-dimensional oriented vector space endowed with a nondegenerate symmetric bilinear form. The Hodge star operator after operating on an algebraic element (vector, tensors) gives the Hodge dual element. For example, in an oriented 3D Euclidean space, an oriented plane can be represented by the exterior product of two basis vectors, and its Hodge dual is the normal vector given by their cross product; conversely, any vector is dual to the oriented plane perpendicular to it, endowed with a suitable bivector. Generalizing this to an n-dimensional vector space, the Hodge star is a one-to-one mapping of k-vectors to (n - k)-vectors; the dimensions of these spaces are the binomial coefficients $\binom{n}{k} = \binom{n}{n-k}$.

Let $\lambda \, \epsilon \bigwedge^p$ then we have a natural linear map from $\bigwedge^{n-p}$ to $\bigwedge^n$ $\mathbf{V}$ given by,

$$\mu \to \lambda \bigwedge \mu \tag{24}$$

But $\bigwedge^n \mathbf{V}$ is a one-dimensional vector space spanned by some element $\sigma$, so

$$\lambda \bigwedge \mu = f_\lambda(\mu)\sigma \tag{25}$$

for some linear functional $f_\lambda$ on $\bigwedge^{n-p}$. Given an inner product g (n-dimensional oriented space with nondegenerate symmetric bilinear form $< ., . >$), the **Riesz lemma** guarantees the existence of a unique (n - p)-vector $\star\lambda$ such that

$$g(\star\lambda, \mu) = f_\lambda(\mu) \tag{26}$$

The element $\star\lambda \ \epsilon \bigwedge^{n-p}$ is called the **Hough dual** or **Hough star** of $\lambda$. And thus,

$$\lambda \bigwedge \mu = g(\star\lambda, \mu)\sigma \tag{27}$$

# 3 MANIFOLDS

## 3.1 Topology

A topology $\tau$ on a set X is a family of subsets of X, called open sets, satisfying the following.

1. Arbitrary unions of open sets are open.
2. Finite intersections of open sets are open.
3. The empty set $\emptyset$ and X are both open.

A **topological space** (or simply a space) is a set X endowed with a topology. Let X be a finite set and $\tau$ be the set of all subsets of X, then $\tau$ is called the **discrete topology** on X. From the point of view of neighborhood, topology is also called the study of **nearness** relation.

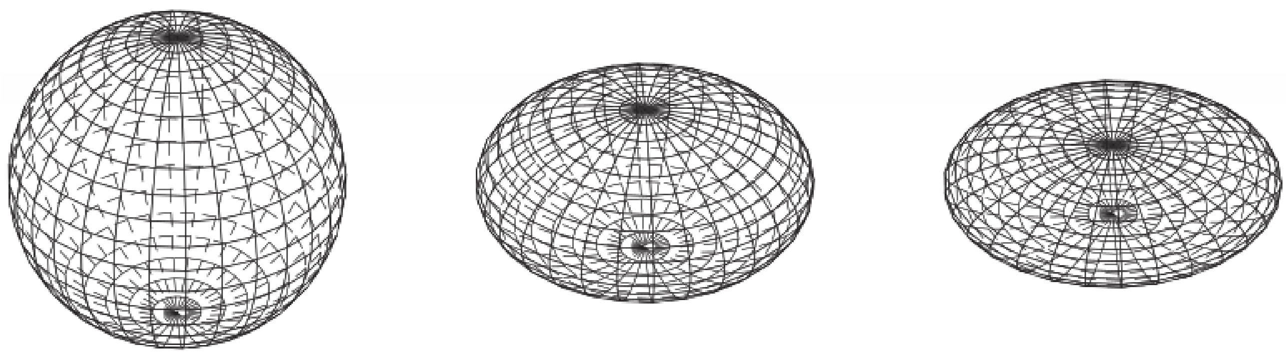

Figure 1: Topological 2-spheres

All these shapes have the same topology but, since the distance between the points on the surface has changed, they have different geometries.

## 3.2 Some Basic Set Terminologies

**a. Open and Closed Sets:** A subset S of X is said to be an **open set** if every point of S is an interior point of S i.e.

$$\forall \text{ x } \epsilon \text{ S } \exists \epsilon > 0 \text{ s.t. } \mathrm{B}(\mathrm{x}; \epsilon) \subset \mathrm{S}.$$

A subset S of X is called a **closed set** if $\overline{\mathrm{S}}$ = X-S is open.

Some sets may be both open and closed, such sets are called **clopen sets**.

**b. Interior and Exterior:** Let S $\subset \mathbb{R}$ and x $\epsilon$ S then x is said to be an **interior point** of S if $\exists$ B(x; r) s.t. B(x; r) $\subset$ S. The set of all interior points of S is called the **interior** of S, **int(S)**. Consequently, the union of all open sets contained in S is the int(S).

Similarly, x is called an **exterior point** of S if x is an interior point of X-S and the set of all exterior points of S is called the **exterior** of S, **ext(S)**.

**c. Closure:** Closure of S $\subset$ X, **Cl(S)** is the intersections of all the closed sets containing S. Set S is called **dense** if Cl(S) = X.

**d. Cover:** An **open cover** of S $\subset$ X is a collection of $U_\alpha$ of open sets in X whose union contains S i.e.

$$S \subseteq U_\alpha$$

An open cover is **locally finite** if, for every x $\epsilon$ S , there is an open neighborhood B(x; r) of x such that $|\{\alpha\text{: } U \cap U_\alpha \neq \emptyset\}|$ is finite. (Note that a locally finite open cover is not necessarily finite.) An open cover $V_\beta$ is a **refinement** of $U_\alpha$ if for all $\beta$ there is an $\alpha$ such that $U_\alpha \supset V_\beta$ .

**e. Compact:** A topological space X is **compact** if every open cover has a finite subcollection that also covers X ("every open cover has a finite **subcover**"). Intuitively, compact spaces can be thought of as being finite in extent. For example, a subset of Euclidean space is compact if and only if it is closed and bounded (**Heine-Borel theorem**).

A **base/basis** for a topology $\tau$ of a topological space (X, $\tau$) is a family/collection, $\mathcal{B}$ of open sets of X s.t. every open set of the topology is equal to the union of some sub-family of $\mathcal{B}$ ($\mathcal{B}$ spans $\tau$). Eg: the set of all open intervals in real number line $\mathbb{R}$ is a basis for Euclidean topology on $\mathbb{R}$ because every open interval is an open set, and every open subset of $\mathbb{R}$ can be written as a union of some family of open intervals. The topology generated by $\mathcal{B}$ is just the collection of open subsets of X where U is considered open if $\forall$ x $\epsilon$ U $\exists$ B $\epsilon\,\mathcal{B}$ s.t. x $\epsilon$ B $\subset$ U.

Given two sets X and Y , the **product topology** on X $\times$ Y is the topology generated by all sets of the form U $\times$ V, where U is open in X and V is open in Y. For Eg: The standard topology on $\mathbb{R}^n$ is the product topology on $\underbrace{\mathbb{R} \times ... \times \mathbb{R}}_{n\ times}$.

## 3.3 Mapping Between Topologies

Let f: X $\rightarrow$ Y be a map between topological spaces. The map f is **continuous** if the inverse image of an open set S in Y is open in X.

Let $\tau$ be a topology on X, and let Y $\subset$ X. The collection $\tau_Y = Y \cap U\text{: } U \,\epsilon\, \tau$ is a topology on Y, called the **induced topology** or **subspace topology** on Y. The space Y equipped with this topology is a **subspace** of X.

A map f: X $\rightarrow$ Y between topological spaces is a **homeomorphism** if it is continuous with continuous inverse, meaning that there is a continuous map g: Y $\rightarrow$ X such that g $\circ$ f = f $\circ$ g = 1. If such a pair of maps exist, we write X $\approx$ Y and say that X and Y are **homeomorphic** (or **topologically equivalent**). A **property** P(X) is a **topological invariant** of X if X $\approx$ Y implies P(X) = P(Y) i.e. P(X) depends only on the topology of X. The property P(X) is a **complete topological invariant** of a space X provided that P(X) = P(Y) iff X $\approx$ Y. Topology can be loosely characterized as the study of the topological invariants of spaces.

## 3.4 Multivariate Calculus

Let U $\subset \mathbb{R}^n$ be an open set and suppose that f: U $\to \mathbb{R}$ is a function. Label the points of $\mathbb{R}^n$ by the n-tuples x = (x$^1$, ..., x$^n$ ). Then the partial derivative $\partial$f /$\partial$x$^i$ is defined by

$$\frac{\partial f}{\partial x^i} = \lim_{h\to 0} \frac{f(x + he_i) - f(x)}{h} \tag{28}$$

where e$_i$ = (0, ..., 1, ..., 0) has a "1" in the ith slot. For higher order, let $\alpha$ = (i$_1$, ..., i$_k$). Then

$$\frac{\partial f}{\partial x^\alpha}(x) = \frac{\partial^k f}{\partial x^{i_1} ... \partial x^{i_k}} \tag{29}$$

A function f: $\mathbb{R}^n \to \mathbb{R}$ is $\mathbf{C}_\infty$, or **smooth**, if $\frac{\partial f}{\partial x^\alpha}$ exists and is continuous for all $\alpha$. The composition of smooth functions is smooth.

Let U $\subset \mathbb{R}^n$ be an open set but now suppose that f: U $\to \mathbb{R}^m$ is a map, given by x $\to$ (f$^1$(x), ..., f$^m$(x)). The map f is smooth if each component function f$^1$ is smooth. The derivative D f(x) of f at x is just the matrix of partial derivatives D f(x) = $(\partial f^i/\partial x^j)$; this matrix is called the **Jacobian matrix**. When n = m, its determinant det$(\partial f^i/\partial x^j)$ is called the Jacobian determinant or more simply the **Jacobian** of the map f.

Let U, V $\subset \mathbb{R}^n$ be two open sets, and let f: U $\to$ V be a homeomorphism. If f and f$^{-1}$ are both smooth then f is called a **diffeomorphism** (isomorphism of smooth manifolds). Since every diffeomorphism is a homeomorphism, given a pair of manifolds which are diffeomorphic to each other they are in particular homeomorphic to each other. The converse is not true in general. This brings us to the all-important **inverse function theorem (IFT)** which states, "if the matrix representing the derivative of a function is invertible at some point then the function itself is a local diffeomorphism in the neighborhood of that point".

## 3.5 Coordinate Systems

2-sphere

Consider the transformation in $\mathbb{R}^2$ between Cartesian and polar coordinates:

$$\begin{aligned} x &= r cos\theta \\ y &= r sin\theta \end{aligned} \tag{30}$$

Every point in the plane can be described uniquely in the (x, y) coordinate system, but the **origin** is a problem for the polar coordinate system because it is described by the infinity of pairs (r, $\theta$) = (0, anything). Another way to see that something strange happens at the origin is to compute the Jacobian of the transformation.

$$\begin{vmatrix} \partial x/\partial r & \partial x/\partial\theta \\ \partial y/\partial r & \partial y/\partial\theta \end{vmatrix} = \begin{vmatrix} cos\theta & -r sin\theta \\ sin\theta & r cos\theta \end{vmatrix} = r \tag{31}$$

Now, at origin r vanishes and consistently the Jacobian determinant.

The **IFT** provides the link between these two ways of determining the validity of a given coordinate transformation. A coordinate transformation is a good one if there is a one-to-one correspondence between the two sets of coordinates and if the transformation is differentiable. In other words, a set of functions $f^i(x)$ on $\mathbb{R}^n$ constitutes a **good coordinate system** in the neighborhood of a point x if the transformation $(x^1, ..., x^n) \to (f^1(x^1, ..., x^n), ..., f^n(x^1, ..., x^n))$ is a diffeomorphism. Ascertaining whether a map is a diffeomorphism can be difficult. But checking whether a Jacobian vanishes is usually easy. The **IFT** assures that, as long as the Jacobian of the transformation is non-singular, we can coordinatize the neighborhood of x with the functions $f^i$.

## 3.6 Differential Manifold (Smooth Manifold or $C_\infty$ )

An n-dimensional smooth manifold **M** consists of a **Hausdorff** topological space together with a **countable** collection of open sets $\{U_i\}$, called **coordinate neighborhoods/patches**, that **cover M** and a collection of maps $\{\psi_i\}$, called **coordinate maps**, satisfying two conditions.

1. Each $\psi_i$: U $\to \mathbb{R}^n$ is a homeomorphism onto an open subset of Rn . (We say that **M** is **locally Euclidean**.)
2. If $U_i$ and $U_j$ are two overlapping coordinate neighborhoods with coordinate maps $\psi_i$ and $\psi_j$ then $\psi_j \circ \psi_i^{-1}$: $\psi_i(U_i \cap U_j) \to \psi_j(U_i \cap U_j)$ is a diffeomorphism. (We say that the coordinate maps are **compatible** on overlaps.)

Each pair $(U_i, \psi_i)$ is called a **coordinate chart**, and the collection of all coordinate charts is called an **atlas**. The maps $\psi_j \circ \psi_i^{-1}$ are called **transition functions** of the atlas.

The **Hausdorff condition** is there basically to express our intuition of space as "infinitely divisible", so that we can separate points with open sets. The condition that the **cover be countable** is there for a technical reason having to do with extending locally defined quantities to globally defined ones. The condition that **M** be **locally Euclidean** serves at least two purposes. First, it tells us that, in the neighborhood of a point, all n-dimensional manifolds look like a (mildly deformed) bit of Euclidean n-space. Second, it allows us to define local coordinates so that we can compute things. The compatibility condition ensures that we can patch together the coordinate systems consistently, so that we always end up with valid coordinates.

Let $(U, \psi)$ be a coordinate chart with $p \in U$ and suppose that $\psi(p) = q$. If $x^1, \dots, x^n$ are the standard coordinate functions on $\mathbb{R}^n$ then q has coordinates $(x^1(q), \dots, x^n(q))$. Thus we can write

$$\psi(p) = (x^1(q), ..., x^n(q)) \tag{32}$$

The functions $x^1, \dots, x^n$, viewed as functions on U , are called **local coordinates** on U. Let V be another coordinate neighborhood, with local coordinates $y^1, \dots, y^n$. If $U \cap V = \phi$ then on $\psi(U \cap V)$ the action of the transition function $\psi_1 \circ \psi_2^{-1}$ can be written in local coordinates as follows:

$$(x^1, ..., x^n) \to (y^1(x^1, ..., x^n), ..., y^n(x^1, ..., x^n)) \tag{33}$$

The compatibility condition (2) in the definition of a manifold is just the statement that equation

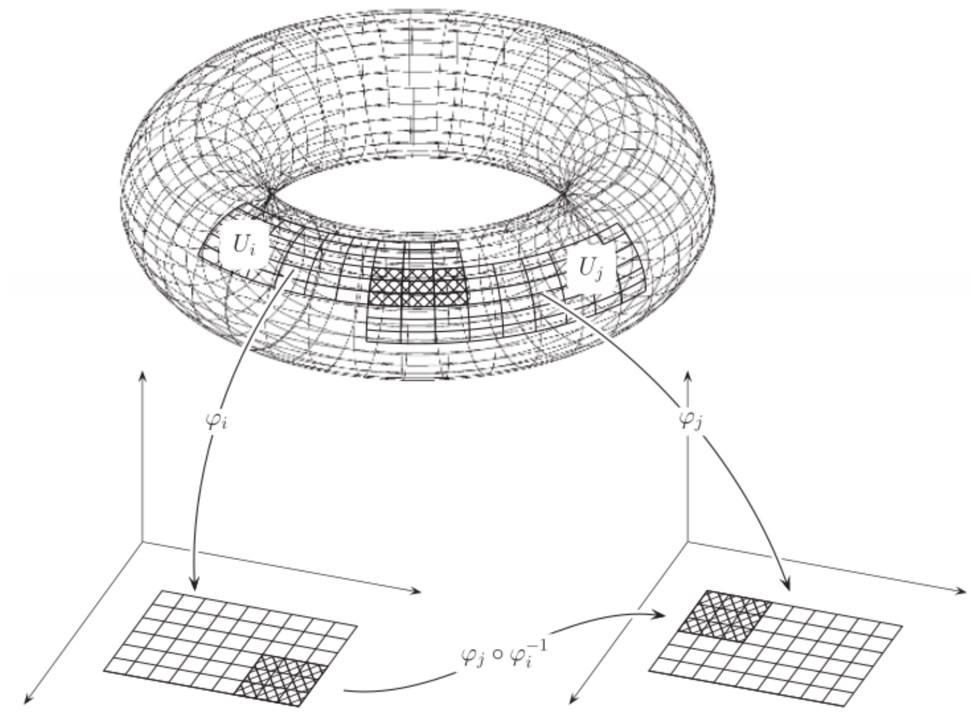


Figure 2: Two coordinate charts and a transition function

33 is a diffeomorphism, which, by the inverse function theorem, is equivalent to the requirement that the Jacobian determinant det($\partial$y$_i$/$\partial$x$_j$) be nonzero. The sign of this determinant is important. A manifold is said to be **orientable** if it is possible to choose an ordering of the local coordinates so that the Jacobian determinants of the transition functions have the same sign on every pair of overlapping neighborhoods. If this is possible then the manifold has two opposite orientations, according to the choice of sign.

## 3.7 Smooth Maps on Manifolds

The definition of the derivative given in Section 3.4 uses the linear structure of Euclidean space in a crucial way, and there is no way to define something similar for a general curved space. Instead, the existence of the coordinate maps on a smooth manifold is used to define differentiability.

A function f: **M** $\to \mathbb{R}$ is smooth at p $\epsilon$ **M** if, for any chart (U, $\psi$) with p $\epsilon$ U, the map

$$\widetilde{f} = f \circ \psi^{-1} : \psi(U) \to \mathbb{R} \tag{34}$$

is a smooth function in the usual Euclidean sense near $\psi$(p).

In general, Let **M** and **N** be two smooth manifolds of dimensions m and n, respectively. A map f: **M** $\to$ **N** is smooth if $\forall$ p $\epsilon$ **M** $\exists$ charts (U, $\psi_1$) on **M** and (V, $\psi_2$) on **N**, with p $\epsilon$ U and f(U)

$\subset$ V s.t.

$$\widetilde{f} = \psi_2 \circ f \circ \psi_1^{-1} : \psi_1(U) \to \psi_2(V) \tag{35}$$

is a smooth map of Euclidean spaces. A smooth map f: $\mathbf{M} \to \mathbf{N}$ is a **diffeomorphism** if $f^{-1}$ exists and is smooth, in which case we say that **M** and **N** are diffeomorphic.

Because smooth maps of manifolds reduce to smooth maps of Euclidean space, the inverse function theorem works for manifolds in the same way as it does in Euclidean space: if the Jacobian matrix of f is non-singular then f is a local diffeomorphism. In particular, M and N can only be **diffeomorphic** if m = n.

## 3.8 Immersion and Embedding

A prototypical theme in geometry is the study of "**spaces with structure**", i.e. a set X equipped with some sort of additional geometric structure, such as a topology in the case of topological spaces. It is important how functions between our spaces with structure interact with the structure on those spaces. We concern ourselves only with those functions which "preserve" the structure (e.g. **differentiable maps** between differentiable manifolds, which are maps respecting the differentiable structure).

A differentiable mapping f: $\mathbf{M}^m \to \mathbf{N}^n$ (m = dim **M**, n = dim **N**) of differentiable manifolds **M** and **N** is said to be an immersion if $df_p$: $T_p\mathbf{M} \to T_{f(p)}\mathbf{N}$ is injective $\forall$ p $\epsilon$ **M** where $T_p\mathbf{X}$ denotes the **tangent space** [3.9] of manifold **X** at point p $\epsilon$ **X**.

Equivalently, f is an immersion if its derivative **df** has a constant rank = m i.e. rank(df) = dim **M** = m. The function itself need not be injective, only its derivative **df** must be.

Since df: $T_p\mathbf{M} \to T_{f(p)}\mathbf{N}$ is a linear map between vector spaces and dim $T_p\mathbf{M}$ = dim **M** = m $\forall$ p $\epsilon$ **M**, it follows by linear algebra that if df is an immersion then dim $\mathbf{M} \leq$ dim **N**. We call dim **N** - dim **M** the **codimension** of df.

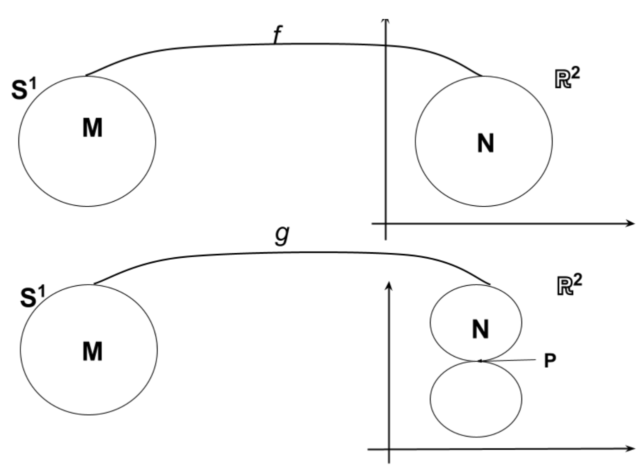


Figure 3: $f$ is an immersion as well as embedding but $g$ is only immersion due to self-intersection at p, which doesn't allow homeomorphism.

If f is an immersion $\forall$ p $\epsilon$ **M** then **M** is an (immersed) **submanifold** of **N**. The Jacobian has **maximal rank** (namely m), so f(M) is locally coordinatizable according to the IFT. If m $\geq$ n and the Jacobian of the transformation has maximal rank (namely n) at p $\epsilon$ **M** then f is called a **submersion** at p.

An injective immersion f is called an **embedding** provided that f maps **M** homeomorphically onto its image f(**M**) (in the induced topology). The basic difference between immersions and embeddings is that the image of an immersion can have **self-intersections** whereas the image of an embedding cannot.

The **Whitney embedding theorem** says that any n-dimensional topological manifold can be embedded in $\mathbb{R}^{2n+1}$, and any n-dimensional smooth manifold can be embedded in $\mathbb{R}^{2n}$. The Möbius strip is a smooth 1-manifold and can be embedded in $\mathbb{R}^3$ and the Klein bottle is a smooth 2-manifold and thus can be embedded in $\mathbb{R}^4$. If you try to construct the Klein bottle in $\mathbb{R}^3$, you will observe that you cannot do it without creating self-intersections somewhere. Whitney's result means that, without loss of generality, we could simply treat manifolds as living in a large Euclidean space. The intrinsic description of a manifold also has physical utility. Einstein modeled the universe as a smooth manifold of a certain type.

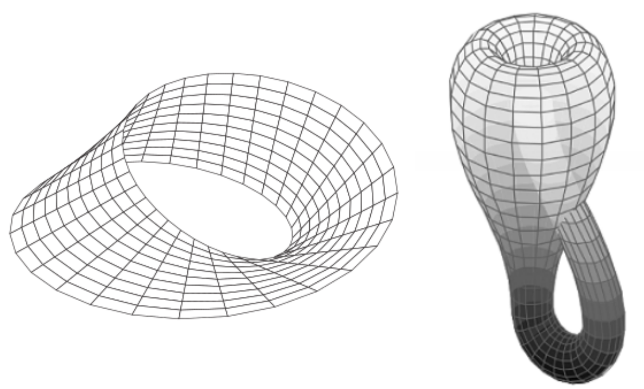

Figure 4: Möbius strip immersed in $\mathbb{R}^2$ (left) and Klein bottle immersed in $\mathbb{R}^3$ (right)

## 3.9 Tangent Space

The **tangent space** is a space spanned by the tangent vectors of curves lying in that space. The tangent space of a manifold generalizes to higher dimensions the notion of tangent planes to surfaces in 3D and tangent lines to curves in 2D. In the context of physics the tangent space to a manifold at a point can be viewed as the space of possible velocities for a particle moving on the manifold ($\overrightarrow{v} = \partial \overrightarrow{r} / \partial t$).

Once the tangent spaces of a manifold have been introduced, one can define vector fields, which are abstractions of the velocity field of particles moving in space. A vector field attaches to every point of the manifold a vector from the tangent space at that point, in a smooth manner. Such a vector field serves to define a generalized ordinary differential equation on a manifold: A solution to such a differential equation is a differentiable curve on the manifold whose derivative at any point is equal to the tangent vector attached to that point by the vector field. All the tangent spaces of a manifold can be glued together to form a new differentiable manifold with twice the dimension of the original manifold, called the **tangent bundle** of the manifold.

Let us define a tangent vector $X_p$ at a point p $\epsilon$ **M** to be a **linear derivation** at p. This means that, $\forall$ a, b $\epsilon$ $\mathbb{F}$ and f, g $\epsilon$ $\Omega^0(\mathbf{M})$, $X_p : \Omega^0(\mathbf{M}) \to \mathbb{R}$ satisfies

1. **linearity**: $X_p$ (af + bg) = a $X_p$ (f) + b$X_p$ (g), and
2. **Leibniz property**: $X_p$ (f g) = g(p)$X_p$ (f) + f (p)$X_p$ (g).

The vector space $T_p$**M** generated by all the $X_p$ is called the **tangent space to M at p**. It is important to note that the tangent space $T_p$**M** is defined independently of any coordinate system. Thus, at each point we are free to pick any basis $\{e_i\}$, not just a coordinate basis $\{\partial/\partial$ $x_i\}$. Any smoothly varying basis on **M** is called a **frame field**. Generally, frame fields do not exist everywhere on **M**; they clearly exist locally, because $e_i = \partial/\partial$ $x_i$ is an instance. In terms of a frame field $\{e_i\}$, we could write a vector field X as

$$X = X_i e_i \tag{36}$$

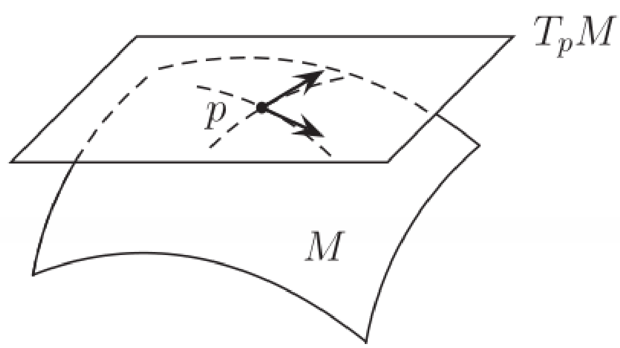


Figure 5: Although the tangent space $T_p$**M** is an abstract linear space attached to **M** at p, we often imagine it to look like above figure.

## 3.10 Cotangent Space

The cotangent space $T_p^*$**M** is defined as the dual space of the tangent space at p, $T_p$**M**. The elements of the cotangent space are called **cotangent vectors** or **tangent covectors**. All cotangent spaces at points on a connected manifold have the same dimension, equal to the dimension of the manifold. All the cotangent spaces of a manifold can be "glued together" (i.e. unioned and endowed with a topology) to form a new differentiable manifold of twice the dimension, the cotangent bundle of the manifold.

Let **M** be a smooth manifold and let p be a point in **M**. Let $T_p$**M** be the tangent space at p. Then the cotangent space at p is defined as the dual space of $T_p$**M**:

$$T_p^*\mathbf{M} = (T_p\mathbf{M})^* \tag{37}$$

Concretely, elements of the cotangent space are **linear functionals** on $T_p\mathbf{M}$ i.e. every element $\alpha \epsilon T_p\mathbf{M}$ is a linear map $\alpha : T_p\mathbf{M} \rightarrow \mathbb{F}$

where $\mathbb{F}$ is the underlying field of the vector space being considered, for example, the field of real numbers. The elements of $T_p^*\mathbf{M}$ are called cotangent vectors.

Let $\{x^i\}$ be local coordinates around p, then $T_p\mathbf{M}$ is spanned by the n basis vectors $\partial/\partial x^i$. The corresponding dual basis vectors are denoted $dx^i$. By definition we have

$$\left\langle \frac{\partial}{\partial x_i}, dx^j \right\rangle = \delta_i^j \tag{38}$$

A general element $\alpha$ p of $T_p^*\mathbf{M}$ is a linear combination of the basis elements:

$$\alpha_p = a_i dx^i \tag{39}$$

where $\alpha_i$ are constants.

A **differential 1-form** or **smooth covector field** on **M** is a smooth map p$\rightarrow$ $\alpha_p$. In local coordinates around p

$$\alpha = a_i(x) dx^i \tag{40}$$

where the $a_i(x)$ are smooth functions on **M**.

**1-forms** (covector/pseudovector) of a vector space with local coordinates ($dx_1$, $dx_2$, ..., $dx_n$) are the elements of a vector space with local basis ($dx^1$, $dx^2$, ..., $dx^n$) i.e. dual space. (Vectors (covariant vectors/**kets** $|\psi\rangle$) and 1-forms (contravariant vectors/**bras** $\langle\psi|$) are dual to each other.) **2-forms** are the elements of the exterior product space with local basis ($dx^1 \bigwedge dx^2$, $dx^2 \bigwedge dx^3$ ..., $dx^{n-1} \bigwedge dx^n$). In general, (**k-form**) (or a form of degree k) are the elements of the exterior product space with basis ($dx^1 \bigwedge dx^2 ... dx^p$, ...). For eg: in 3D Euclidean space with basis (i, j, k),

1-form have basis (i', j', k')
2-form have basis (j$\bigwedge$k, k$\bigwedge$i, i$\bigwedge$j)
3-form have basis (i$\bigwedge$j$\bigwedge$k)
other higher forms don't exist for 3D.

Like tangent space, cotangent space is also independent of any particular coordinate system for its definition. If $\{e_i\}$ is a generic basis for $T_p\mathbf{M}$ then the corresponding dual basis $\{\theta^i\}$ is called a **coframe field**, so

$$\left\langle e_i, \theta^j \right\rangle = \delta_i^j \tag{41}$$

## 3.11 Cotangent Space as Jet Space

We have defined a differential 1-form as a smoothly varying map on M that assigns to every point p of M an element of the cotangent space at p. Although this is correct, it is perhaps rather unsatisfying because the cotangent space is defined in terms of the tangent space, which is itself defined as a space of derivations on functions. There is, however, an elegant way to define the cotangent space directly, using just the notion of a smooth function, which is perhaps a bit more natural from the point of view of manifold theory. It uses the idea of a **jet**. Jet is an operation that takes a differential function $f$ and produces a polynomial, the truncated Taylor polynomial of $f$ at each point of its domain. The theory of jets regards these polynomials as abstract polynomials rather than polynomial functions.

$$(J^k_{x_0} f)(z) = \sum_{i=0}^{k} \frac{f^{(i)}(x_0)}{i!} z^i = f(x_0) + f'(x_0)z + ... + \frac{f^{(k)}(x_0)}{k!} z^k \tag{42}$$

gives the k-jet of f:$\mathbb{R} \rightarrow \mathbb{R}$ at point $x_0$ in 1D case.

Let $C^\infty(\mathbb{R}^n, \mathbb{R}^m)$ be a vector space of smooth functions f:$\mathbb{R}^n \rightarrow \mathbb{R}^m$. Let k is a non-negative integer and p $\epsilon \mathbb{R}^n$, then we can define an equivalence relation $\mathbf{E}^k_p$ on $C^\infty$ by declaring that two functions $f$ and $g$ are equivalent to the order of k if they have the same value at p, and all their partial derivatives agree at p upto their kth derivatives i.e. $f \sim g$ iff $f$-$g$ = 0 to kth order. The **kth order jet space** of $C^\infty(\mathbb{R}^n, \mathbb{R}^m)$ at p is defined as the set of equivalence classes of $\mathbf{E}^k_p$ and is denoted by $J^k_p(\mathbb{R}^n, \mathbb{R}^m)$.

Now, let $f$:$\mathbf{M} \rightarrow \mathbb{R}$ be a smooth function, p$\epsilon\mathbf{M}$, and $\{x^i\}$ local coordinates around p. We say that $f$ vanishes to first order at p if $\partial f/\partial x^i$ vanishes at p for all i. for k$\geq$1 we say that $f$ vanishes to kth order at p if, $\forall$ i, $\partial f/\partial x^i$ vanishes to (k-1)th order at p, where $f$ vanishes to zeroth order at p if $f$(p) = 0. In other words, $f$ vanishes to kth order at p if the first k terms in its Taylor expansion vanish at p.

For k>0 let $\mathbf{M}^k_p$ be the set of all smooth functions on $\mathbf{M}$ vanishing to (k-1)th order at p, and set $\mathbf{M}^0_p := \Omega^0(\mathbf{M})$ and $M_p := \mathbf{M}^1_p$. Each $\mathbf{M}^k_p$ is a vector space under the usual pointwise operations, and we have the series of inclusions

$$\mathbf{M}^0_p \supset \mathbf{M}^1_p \supset \mathbf{M}^2_p \supset ....$$

We now define $T^*_p\mathbf{M}$, the cotangent space to $\mathbf{M}$ at p, to be the **quotient space**

$$T^*_p\mathbf{M} = \mathbf{M}_p/\mathbf{M}^2_p \tag{43}$$

An element of $T^*_p\mathbf{M}$ is called a **differential 1-form** at p.

## 3.12 Tensor Field

There are many different ways to define a tensor field on a manifold, but the essential idea is simple enough: just like a vector field is an assignment of vectors to each point in space (manifold $\mathbf{M}$), a tensor field is just a smooth assignment of a tensor to every point of $\mathbf{M}$. A tensor field $\Psi$ of type (r, s) can also be defined as a map from $\mathbf{M}$ to $\mathbf{T}^s_r$ s.t. on any coordinate patch U, the components $\Psi^{i_1 i_2 ... i_r}_{j_1 j_2 ... j_s}$ are smoothly varying functions.

$$\Psi = \Psi^{i_1 i_2 ... i_r}_{j_1 j_2 ... j_s} \frac{\partial}{\partial x^{i_1}} \otimes ... \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes ... \otimes dx^{j_s} \tag{44}$$

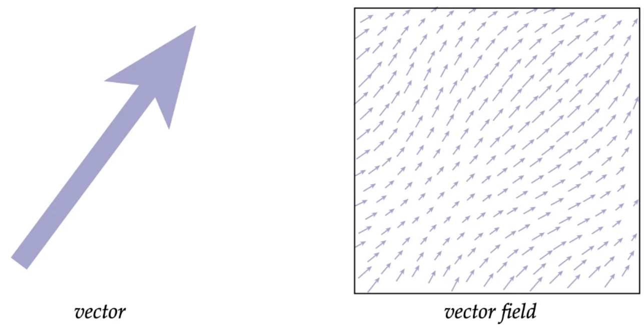


Figure 6: Just like a vector field is an assignment of vectors to each point in space (manifold), a tensor field is an assignment of tensors to each point in space (manifold)

where for any p $\epsilon$ U , the vector fields $\frac{\partial}{\partial x^{i_i}}$ and the 1-form fields $dx^i$ constitute dual bases for $T_p\mathbf{M}$ and $T_p^*\mathbf{M}$ respectively, thus forming a basis of tensor product space $(T_p\mathbf{M})^{\otimes r} \otimes (T_p^*\mathbf{M})^{\otimes s}$.

Tensor field can also be viewed as a smoothly varying multilinear map. Let $\widetilde{\mathbf{T}}_r^s(p)$ be the space of all multilinear maps on $(T_p^*\mathbf{M})^{\times r} \times (T_p\mathbf{M})^{\times s}$. A tensor field $\Psi$ of type (r, s) is then a smooth assignment of an element $\Psi_p \,\epsilon\, \widetilde{\mathbf{T}}_r^s(p)$ to each p $\epsilon$ $\mathbf{M}$. The map $\Psi$ must be more than just multilinear-in fact, it must be functional multilinear, which means that equation 45 must hold for $\Psi$.

$$\Psi(v_1, ..., \alpha_1 u + \alpha_2 w, ..., v_{r+s}) = \alpha_1 \Psi(v_1, ..., u, ..., v_{r+s}) + \alpha_2 \Psi(v_1, ..., w, ..., v_{r+s}) \tag{45}$$

where $\alpha_1$ and $\alpha_2$ are smooth functions on $\mathbf{M}$. And thus,

$$\Psi : \underbrace{\Gamma(T^*\mathbf{M}) \times ... \times \Gamma(T^*\mathbf{M})}_{r\ times} \times \underbrace{\Gamma(T\mathbf{M}) \times ... \times \Gamma(T\mathbf{M})}_{s\ times} \to \Omega^0(\mathbf{M}) \tag{46}$$

where $\Gamma(T\ \mathbf{M})$ and $\Gamma(T^*\ \mathbf{M})$ denote the space of all vectors and covectors fields on $\mathbf{M}$.

## 3.13 Differential Forms

Just as a k-vector is a special kind of tensor (alternating tensor), a k-form is a special kind of tensor field (alternating tensor field). k-forms are by far the most useful kind of tensor fields, for several reasons:

1.They are easy to define.
2. They are easy to use because, for the most part, one does not have to deal with all those irritating indices.
3. Almost all important geometrical quantities can be expressed in terms of forms.
4. Differential forms are essentially the things that appear under integral signs.

Let U be a coordinate patch of $\mathbf{M}$ and let p $\epsilon$ U. A k-form $\omega_p$ on U at p is an element of k $(T_p^*\mathbf{M})$. It follows that, in local coordinates,

$$\omega_p = \frac{1}{k!} \sum a_{i_1 ... i_k} dx^{i_1} \bigwedge ... \bigwedge dx^{i_k} = \sum a_I dx^I \tag{47}$$

for some constants $a_I = a_{i_1...i_k}$. A (differential) k-form $\omega$ on **M** is a smooth assignment p $\rightarrow \omega_p$. In local coordinates,

$$\omega_U = \frac{1}{k!}\sum a_{i_1...i_k}(x)dx^{i_1}\bigwedge...\bigwedge dx^{i_k} = \sum a_I(x)dx^I \tag{48}$$

where the $a_I$(x) are now smooth functions on U. The vector space of all k-forms on **M** is denoted $\Omega^k$(**M**). In particular, we think of the smooth functions on **M** as **0-forms** as shown in equation 46.

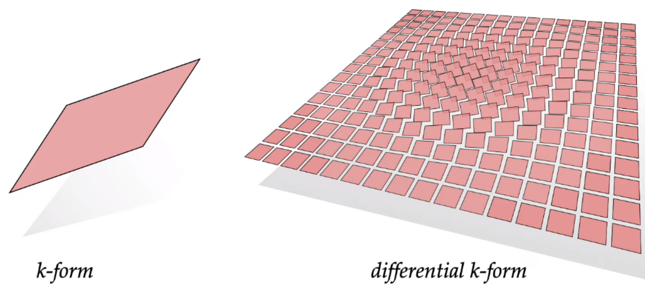


Figure 7: Just like a vector field is an assignment of vectors to each point in space (manifold), a differential k-form is an assignment of k-forms to each point in space (manifold)

### 3.14 Exterior Derivative

We now introduce a natural differential operator on k-forms, denoted by 'd'. It is a far reaching generalization of the ordinary gradient, curl, and divergence operators of multivariate calculus.

**Theorem:** There exists a unique linear operator d:$\Omega^k$(**M**) $\rightarrow$ $\Omega^{k+1}$(**M**) called the exterior derivative, satisfying the following properties. For any forms $\lambda$ and $\mu$, and for any function $f$ , the operator d is

1. **linear:** d($\lambda$ + $\mu$) = d$\lambda$ + d$\mu$,
2. **graded derivation:** d($\lambda \bigwedge \mu$) = d$\lambda \bigwedge \mu$ + (-1)$^{(deg\lambda)}\lambda\bigwedge$ d$\mu$,
3. **nilpotent:** d$^2\lambda$ = 0, and
4. **natural:** in local coordinates {$x_i$} about a point p, d$f$ = $\sum \frac{\partial f}{\partial x^i}dx^i$.

### 3.15 Interior Product (Interior Derivative)

Interior product is a degree -1 (anti)derivation on the exterior algebra of differential forms on a smooth manifold. Given a vector field X we can define a linear map $i_X$ : $\Omega^k$(**M**) $\rightarrow \Omega^{k-1}$ (**M**), taking k-forms to (k-1)-forms, called the interior product, satisfying the following properties. Let f be a function, $\omega$ a 1-form, and $\lambda$ and $\eta$ arbitrary-degree forms. Then

1. $i_X$f = 0
2. $i_X\omega = \omega$(X) := $\langle \omega, X\rangle$
3. **graded derivation:** $i_X(\lambda\bigwedge\eta)$ = $i_X\lambda\bigwedge\eta$ + (-1)$^{deg\lambda}\lambda\bigwedge$ $i_X\eta$.

## 3.16 Pullback and Pushforward

Let $f$: $\mathbf{M}\rightarrow\mathbf{N}$ be a smooth map of manifolds. Given a smooth map $g$: $\mathbf{N}\rightarrow\mathbb{R}$ we define a new map $f^{*}g$: $\mathbf{M}\rightarrow\mathbb{R}$ by $f^{*}g = g\circ f$. The map $f^{*}g$ is called the **pullback** of $g$ by $f$, because the function $g$ is "pulled back" from $\mathbf{N}$ to $\mathbf{M}$. In simple terms, pullback is the composition of smooth maps. A pullback map is extended to forms by requiring it to be

1. **a ring homomorphism:** $f^{*}(\lambda+\mu) = f^{*}\lambda + f^{*}\mu$, $f^{*}(\lambda\bigwedge\mu) = f^{*}\lambda\bigwedge f^{*}\mu$ and
2. **natural:** $\mathrm{d}(f^{*}\lambda) = f^{*}(\mathrm{d}\lambda)$.

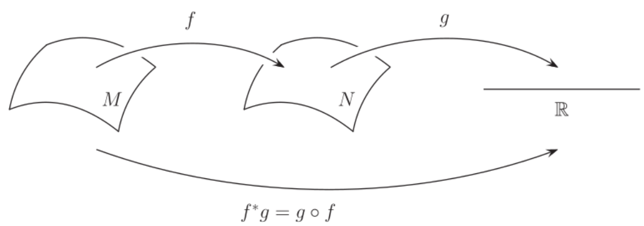


Figure 8: Pullback map $g\circ f$ from $\mathbf{N}$ to $\mathbf{M}$

Pushforward and Pullback are **dual** to each other. The pullback map $f^{*}$ naturally pulls a form on $\mathbf{N}$ back to a form on $\mathbf{M}$. One can ask a similar question regarding vector fields. Do they pull back as well? For that we introduce a new linear map $f_{*}$: $\mathrm{T}_{p}\mathbf{M}\rightarrow \mathrm{T}_{f(p)}\mathbf{N}$ of tangent spaces, called the **pushforward** by $(f_{*}X_{p})_{f(p)}(g) := X_{p}(f^{*}g)$. It is a linear approximation of smooth maps of manifolds on tangent spaces. Pushforward is also called differential of $f$ as it is generalization of the derivative in calculus. It can be used to push tangent vectors on $\mathbf{M}$ forward to tangent vectors on $\mathbf{N}$.

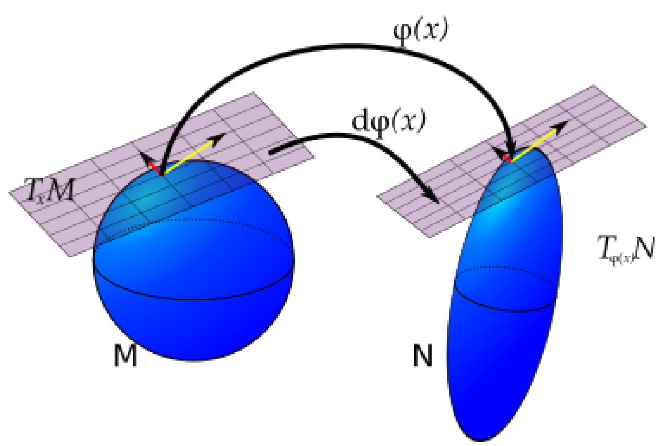


Figure 9: If a map $\phi$, carries every point on manifold $\mathbf{M}$ to manifold $\mathbf{N}$ then the pushforward of $\phi$ carries vectors in the tangent space at every point in $\mathbf{M}$ to a tangent space at every point in $\mathbf{N}$.

In terms of the pushforward map, the **inverse function theorem** can be stated as follows: the map $f$: $\mathbf{M}\rightarrow\mathbf{N}$ is a local diffeomorphism in the neighborhood of a point p iff $f_{*}$ is an isomorphism of $T_p\mathbf{M}$ and $T_{f(p)}\mathbf{N}$. Thus,we can push vectors forward by any smooth map but you can only push vector fields forward by a diffeomorphism.

### 3.17 Integral Curves and Lie Derivative

In addition to the exterior derivative of a differential form there is another intrinsic derivative (its definition does not require the introduction of any additional structure) on manifolds. If X is a vector field on **M** then, X can be thought of as a generalized directional derivative operator on functions. But to be able to take derivatives of tensor fields, we require pushforward maps since X does not appear to act directly on tensor fields.

Given a vector field X on **M**, the fundamental theorem of ODE guarantees that $\exists$ a unique maximal curve $\gamma(t)$: $\mathbf{I}\to\mathbf{M}$ through each point p $\epsilon$ **M** s.t. $\gamma(0) = $ p and the tangent vector to the curve $\gamma$ at $\gamma(t)$ is precisely $X_{\gamma(t)}$. The curve $\gamma$ is called the **integral curve** of X through p.

In ordinary calculus, given a curve $\gamma(t)$ (a map to some Euclidean space) $\gamma'(t)$ is naturally a vector (tangent vector to the curve). But this interpretation is no longer tenable in a general manifold that is not itself a vector space. Instead, vectors are derivations, so we really want to understand $\gamma'(t)$ as a derivation. For this we define **I** as a manifold (open interval on real line) and attach the coordinate $t$ to the points of **I**. Then $d/dt$ is a tangent vector field on **I** which acts on functions $f$: $\mathbf{I} \to \mathbb{R}$ according to $(d/dt)f = df/dt$. Then we can define

$$\gamma'(t) = \gamma_*(d/dt)_{\gamma(t)} \tag{49}$$

This allows us to reduce the question of the existence of an integral curve in a general manifold M to a system of ODEs. The existence of integral curves enables us to define the notion of a flow. At each point p $\epsilon$ **M** we can imagine moving a parameter a distance $t$ along the integral curve $\gamma$ of X through p. If we do this for all points of **M** simultaneously we get a **flow** along X. We can think of this movement as the flow of a fluid as shown in the figure 10.

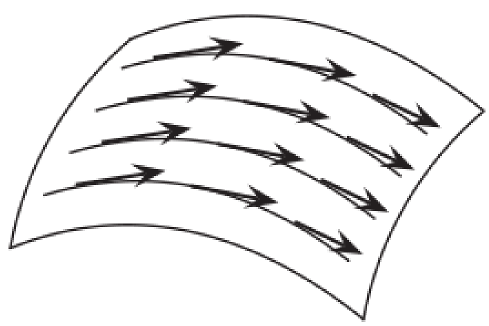

Figure 10: Flowing along the integral curves of a vector field

For any open set **U** of **M** and any p $\epsilon$ **U**, define $\gamma_p(t)$ to be the integral curve of X through p, with p $= \gamma_p(0)$. Then the map $\phi_t$: $\mathbf{U} \to \mathbf{M}$ given by $\gamma_p(0) \to \gamma_p(t)$ is a **flow**, or a **one-parameter group of diffeomorphisms** of **M** generated by the vector field X. Using this flow, we can generalize the notion of the directional derivative to tensor fields. In terms of flow $\phi_t$,

$$(Xf)_p = \lim_{t\to 0} \frac{f_{\phi_t p} - f_p}{t} \tag{50}$$

where $f_p = f(p)$ and $\phi_t p = \gamma_p(t)$

But we cannot do the same with tensors as $\mathrm{T}_p$ and $\mathrm{T}_{\phi_t p}$ lie on different unrelated vector spaces and thus we cannot subtract them. However, if we drag $\mathrm{T}_{\phi_t p}$ back to p (using pullback) by means of flow $\phi_t$ then we can compare its value with $\mathrm{T}_p$. So, the right way to define is,

$$(Xf)_p = \lim_{t\to 0} \frac{\phi_t^* f - f}{t}\Bigg|_p \tag{51}$$

By analogy we can define our generalized directional derivative $\mathcal{L}_X$ (**Lie derivative**) of a tensor field T in the direction X at p by

$$(\mathcal{L}_X T)_p = \lim_{t\to 0} \frac{\phi_t^* T - T}{t}\Bigg|_p = \frac{d}{dt}\phi_t^* T\Bigg|_p \tag{52}$$

**Properties of Lie Derivative:**

1. **preserves tensor type**
2. is **linear:** $\mathcal{L}_X(S+T) = \mathcal{L}_X S\, T + S\, \mathcal{L}_X T$
3. is a **derivation:** $\mathcal{L}_X(S \otimes T) = \mathcal{L}_X \otimes S\, T + S\, \mathcal{L}_X \otimes T$
4. is compatible with the **natural pairing** of forms and vector fields: $\mathcal{L}_X\langle \omega, Y\rangle = \langle \mathcal{L}_X \omega, Y\rangle$ + $\langle \omega, \mathcal{L}_X Y\rangle$
5. **commutation:** For any vector field Y, $\mathcal{L}_X$Y = [X, Y]
6. For vector fields X, Y, and Z, $[\mathcal{L}_X, \mathcal{L}_Y]$Z = $\mathcal{L}_{[X,Y]}$Z

# 4 GROUP THEORY

## 4.1 Group Definition

A **Group** G, is a set with a rule for assigning to every (ordered) pair of elements, a third element, satisfying:

1. If f, g $\epsilon$ G then h = fg $\epsilon$ G.
2. Associativity: $\forall$ f, g, h $\epsilon$ G, f(gh) = (fg)h.
3. Existence of identity element: $\forall$ f $\epsilon$ G $\exists$ $e$ s.t. $e$f = f$e$ = f.
4. Existence of inverse element: $\forall$ f $\epsilon$ G $\exists$ f $^{-1}$ s.t. f f$^{-1}$ = f$^{-1}$f = $e$.

| \ | $e$ | $a$ | $b$ |
|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ |
| $a$ | $a$ | $b$ | $e$ |
| $b$ | $b$ | $e$ | $a$ |

Figure 11: $Z_3$ group multiplication table. (Every row and column of the multiplication table contains each element of the group exactly once. This must be the case because the inverse exists.)

A group G is **finite** if it has a finite number of elements. Otherwise it is **infinite**. The number of elements in a finite group G is called the **order** of G. For eg: $Z_3$, the cyclic group of order 3.

An **Abelian group** G in one in which the multiplication law is commutative i.e. $g_1g_2 = g_2g_1$. And the one which doesn't follows commutation is called **non-Abelian group**.

## 4.2 Representation

A **Representation** of a group G is a mapping D of the elements of G onto a set of linear operators with the following properties:

1. D($e$) = 1, where 1 is the identity operator in the space on which the linear operators act.
2. D($g_1$)D($g_2$) = D($g_1g_2$) i.e. the group multiplication law is mapped onto the natural multiplication in the linear space on which the linear operators act.

For eg: representation of $Z_3$ is,

$$D(e) = 1, \qquad D(a) = e^{2\pi i/3}, \qquad D(b) = e^{4\pi i/3} \tag{53}$$

another representation of $Z_3$ can be directly constructed from the multiplication table as,

$$D(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad D(a) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad D(b) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{54}$$

By taking the group elements themselves to form an orthonormal basis for a vector space, $|e\rangle$, $|a\rangle$, and $|b\rangle$ we can also define **regular representation** as,

$$D(g_1)|g_2\rangle = |g_1 g_2\rangle \tag{55}$$

The **dimension of a representation** is the dimension of the space on which it acts and the dimension of the regular representation is the order of the group. The representation of $Z_3$ is 1 dimensional. For any finite group, we can define a vector space in which the basis vectors are labeled by the group elements. Then equation 55 defines the regular representation.

## 4.3 Subgroup

A group H whose elements are all elements of a group G is called a subgroup of G. The identity, and the group G are trivial subgroups of G. For eg., the permutation group S3, has a Z3 subgroup formed by the elements $\{e, a_1, a_2\}$. Subgroup can be used to divide up the elements of the group into subsets called **cosets**. Given an element $g$ of G, the **left cosets** of H in G are the sets obtained by multiplying each element of H by a fixed element $g$ of G (where $g$ is the **left factor**)

$$g\mathrm{H} = \{g\mathrm{h} : \mathrm{h} \in \mathrm{H}\} \; \forall \; g \in \mathrm{G}.$$

The **right cosets** can be defined similarly where g is now a **right factor**.

$$\mathrm{H}g = \{\mathrm{h}g : \mathrm{h} \in \mathrm{H}\} \; \forall \; g \in \mathrm{G}.$$

The number of elements in each coset is the order of H. Every element of G must belong to one and only one coset. Thus for finite groups, the order of a subgroup H must be a factor of order of G. A subgroup H of G is called an **invariant** or **normal subgroup** if $\forall$ g$\in$G

$$gH = Hg \tag{56}$$

i.e.$\forall$ $g$ $\in$G and $h_1 \in$ H $\exists$ an $h_2 \in$ H s.t.

$$h_1 g = g h_2, \text{ or } g h_2 g^{-1} = h_2.$$

The trivial subgroups e and G are invariant for any group. If H is invariant then H$g_1$ H$g_1^{-1}$ = H, so the product of elements in two cosets is in the coset represented by the product of the elements. In this case, the coset space G/H, is called the **factor group** of G by H.

The **center** of a group G is the set of all elements of G that commute with all elements of G. The center is always an Abelian, invariant subgroup of G. However, it may be trivial, consisting only of the identity, or of the whole group.

The **characters** $\chi_D$(g) of a group representation D are the traces of the linear operators of the representation or their matrix elements:

$$\chi_D(g) \equiv TrD(g) = \sum_i [D(g)]_{ii} \tag{57}$$

The advantage of the characters is that because of the cyclic property of the trace Tr(AB) = Tr(BA), they are unchanged by similarity transformations, thus all equivalent representations have the same characters. The characters are also different for each inequivalent irreducible representation, $D_a$—in fact, they are orthonormal up to an overall factor of N.

## 4.4 Eigenstates

In quantum mechanics, we are often interested in the eigenstates of an invariant hermitian operator, in particular the Hamiltonian, H. We can always take these eigenstates to transform according to irreducible representations of the symmetry group. To prove this, note that we can divide up the Hilbert space into subspaces with different eigenvalues of H. Each subspace furnishes a representation of the symmetry group because D($g$), the group representation on the full Hilbert space, cannot change the H eigenvalue because [D($g$), H] = 0. But then we can completely reduce the representation in each subspace. If some irreducible representation appears only once in the Hilbert space, then the states in that representation must be eigenstates of H (and any other invariant operator). This is true because H$|a, j, x\rangle$ must be in the same irreducible representation, thus

$$H|a, j, x\rangle = \sum_y c_y|a, j, x\rangle \tag{58}$$

and if x andy take only one value, then $|a, j, x\rangle$ is an eigenstate.

**Theorem:** If a hermitian operator H, commutes with all the elements D($g$), of a representation of the group G, then you can choose the eigenstates of H to transform according to irreducible representations of G. If an irreducible representation appears only once in the Hilbert space, every state in the irreducible representation is an eigenstate of H with the same eigenvalue.

For Abelian groups, this procedure of choosing the H eigenstates to transform under irreducible representations is analogous to simultaneously diagonalizing H and D($g$) because for an Abelian group that commutes with the H, the group elements can simultaneously diagonalized along with H. This is a consequence of theorem,

**Theorem:** All of the irreducible representations of a finite Abelian group are 1-dimensional.

For a non-Abelian group, we cannot simultaneously diagonalize all of the D($g$)s, but we can completely reduce the representation on each subspace of constant H. A classical problem which is quite analogous to the problem of diagonalizing the Hamiltonian in quantum mechanics is the problem of finding the normal modes of small oscillations of a mechanical system about a point of stable equilibrium. Here, the square of the angular frequency is the eigenvalue of the $M^{-1}K$ matrix and the normal modes are the eigenvectors of $M^{-1}K$.

## 4.5 Tensor Product Representation

Suppose that $D_1$ is an m-dimensional representation acting on a space with basis vectors $|j\rangle$ for j = 1 to m and $D_2$ is an n-dimensional representation acting on a space with basis vectors $|x\rangle$ for x = 1 to n. We can make an m × n dimensional space called the **tensor product space** by taking basis vectors labeled by both j and x in an ordered pair $|j, x\rangle$. Then when j goes from 1 to m and x goes from 1 to n, the ordered pair (j, x) runs over m × n different combinations. On this product space, we can define a new representation called the **tensor product representation** $D_1 \otimes D_2$ by multiplying the two smaller representations. More precisely, the matrix elements of $D_{D_1 \otimes D_2}(g)$ are products of those of $D_1(g)$ and $D_2(g)$:

$$\langle j, x|D_{D_1 \otimes D_2}(g)|k, y\rangle \equiv \langle j|D_1(g)|k\rangle\langle x|D_2(g)|y\rangle \tag{59}$$

## 4.6 Symmetry Group $S_n$ (Permutation Group)

A permutation group is a group G whose elements are permutations of a given set M and whose group operation is the composition of permutations in G. The group of all permutations of a set M is the symmetric group of M, often written as Sym(M) or $S_n$. The term permutation group thus means a subgroup of the symmetric group $S_n$.

Any element of the permutation group on n objects, called $S_n$. can be written in term of cycles, where a cycle is a cyclic permutation of a subset. Commonly used notation is where each cycle is written as a set of numbers in parentheses, indicating the set of things that are cyclicly permuted. For eg.:

(1) means $x_1 \to x_1$
(1372) means $x_1 \to x_3 \to x_7 \to x_2 \to x_1$
(1372)(4) means $x_1 \to x_3 \to x_7 \to x_2 \to x_1$ while $x_4$ remains unchanged. Thus, (1372)(4) = (1372)

Let a particular permutation of a set M = {1, 2, 3, 4, 5} written as

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \tag{60}$$

This means that $\sigma$ satisfies $\sigma(1) = 2$, $\sigma(2) = 5$, $\sigma(3) = 4$, $\sigma(4) = 3$, and $\sigma(5) = 1$. It can be written in cycle notation as $\sigma = (125)(34)$. An arbitrary element has $k_i$ i-cycles, where

$$\sum_{i=1}^{n} ik_i = n \tag{61}$$

$\sigma$ has one 3-cycle and one 2-cycle, so $k_1 = 1$ and $k_2 = 1$.

## 4.7 Conjugacy Class

Conjugacy in abstract algebra is analogous to similarity transformation in linear algebra which relates the linear transformations behaving in similar fashion under change of basis, Let G be a group and $g_1$, $h$ $\epsilon$ G. We say $g_1$ and $h$ are conjugate, $g_1 \sim h$ if $\exists$ $g$ $\epsilon$ G s.t.

$$h = gg_1g^{-1} \tag{62}$$

This implies conjugacy is an **equivalence relation**. Its equivalence classes are called **conjugacy classes**. The conjugacy classes are just the cycle structure, that is they can be labeled by the integers $k_i$. For example, all interchanges are in the same conjugacy class–it is enough to check that the inner automorphism $gg_1g^{-1}$ doesn't change the cycle structure of $g_1$ when $g$ is an interchange, because we can build up any permutation from interchanges. For eg. for a set M = {1, 2, 3, 4}:

(12)(3)(4)·(1)(23)(4)·(12)(3)(4) (note that an interchange is its own inverse)
$\implies \underbrace{1234 \rightarrow 2134}_{(12)(3)(4)}\underbrace{\rightarrow 3124}_{(1)(23)(4)}\underbrace{\rightarrow 3214}_{(12)(3)(4)}$
$\implies$ (13)(2)(4)

Thus, (13)(2)(4) and (1)(23)(4) are conjugate of each other. And the conjugacy class of $S_4$ is the combination of all possible permutations i.e. for $S_4$

$S_4 =$

| | | | |
|---|---|---|---|
| () | (12) | (13) | (14) |
| (23) | (24) | (34) | (123) |
| (132) | (124) | (142) | (134) |
| (143) | (234) | (243) | (1234) |
| (1243) | (1423) | (1324) | (1342) |
| (1432) | (12)(34) | (13)(24) | (14)(23) |

The permutations with same cycle types are a common conjugacy class. For eg.: in above table of $S_4$,the transpositions (12), (13), (14), (23), (24), and (34) are a conjugacy class (conjugate to one another) and similarly for others. Thus, $S_4$ has 5 conjugacy classes.

**Theorem:** In $S_n$ $g \sim h$ iff $g$ and $h$ have the same cycle type.

## 4.8 Lie Group

### 4.8.1 Definition

Lie group ($\mathbf{L}$, $\bullet$) is a group that is also differentiable manifold. Lie group ($\mathbf{L}$, $\bullet$) is

i. a group with group operator $\bullet$
ii. $\mathbf{L}$ is a smooth manifold and
iii. the maps: (group operation of multiplication) $\mu : \mathbf{L} \times \mathbf{L} \rightarrow \mathbf{L}$ which maps $(g_1, g_2) \rightarrow g_1 \bullet g_2$ and (group operation of inversion) $i : \mathbf{L} \rightarrow \mathbf{L}$ which maps $g \rightarrow g^{-1}$ are both smooth maps.

This means a Lie group is a group with a geometric and algebraic structure and the structure must be compatible in a precise way. For example let us consider a $S^1$ group of complex unit circles defined as,

$$S^1 := \{z \,\epsilon\, \mathbb{C} \mid |z| = 1\} \tag{63}$$

and let the operator $\bullet$ be the multiplication on complex numbers $^*\mathbb{C}$. $S^1$ is obviously a group. For two complex numbers $z_1$ and $z_2$ the multiplication operator $^*\mathbb{C}$ multiplies the radii of and adds the angles as

$$z_3 = z_1 * z_2 = r_1 e^{i\theta_1} * r_2 e^{i\theta_2} = r_1 * r_2 e^{i(\theta_1+\theta_2)} = e^{i\theta_3} \tag{64}$$

$r_1 = r_2 = 1$, since $|z| = 1$.

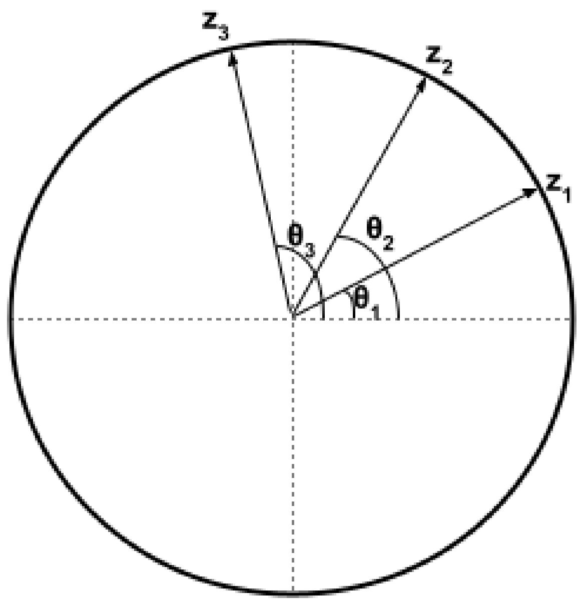


Figure 12: $S_1$ group as Lie group.

This preserves the continuity of symmetry, so $S^1$ is smooth. Also, due to definition of multiplication and inversion, both the group operation of multiplication and inversion are also manifold. This makes the group $S^1$ a Lie group. Moreover, the group operations are commutative, so the Lie group is **Abelian**.

### 4.8.2 Lie Algebra

An algebra, A over a field is a vector space over $\mathbb{C}$ equipped with bilinear product * defined as *:A $\times$ A $\rightarrow$ A. Algebra is defined over a field implies that the vector space is also equipped with addition, multiplication, and scalar multiplication from the field. Typically, algebra follows scalar multiplication and distributivity while commutativity and associativity are optional. Let $\alpha, \beta, \gamma \,\epsilon$ A and c $\epsilon\, \mathbb{C}$ then

i. c $(\alpha * \beta) = (\text{c } \alpha) * \beta = \alpha * (c\,\beta)$
ii. $\alpha * (\beta + \gamma) = (\alpha * \beta) + (\alpha * \gamma)$

Similarly, Lie algebra $\mathcal{L}$ is an algebra with operator [ , ] defined as [ , ]:$\mathcal{L} \times \mathcal{L} \to \mathcal{L}$. Besides general properties of an algebra, Lie algebra has two other properties: $\forall\, \alpha, \beta, \gamma \,\epsilon\, \mathcal{L}$

i. **skew-symmetry:** $[\alpha, \beta]$ = - $[\beta, \alpha]$ and $[\alpha, \alpha] = 0$
ii. **Jacobi identity:** $[\alpha,\, [\beta, \gamma]] + [\beta,\, [\gamma, \alpha]] + [\gamma,\, [\alpha, \beta]] = 0$

These two properties are similar to vector multiplication.

### 4.8.3 Exponential Map

Exponential map is a map from Lie algebra $\mathcal{L}$ of a Lie group **L** to the group, **exp:** $\mathcal{L} \to$ **L**, which allows one to recapture the local group structure from the Lie algebra. The ordinary exponential function is a special case of the exponential map when group G is the multiplicative group of positive real numbers (whose Lie algebra is the additive group of all real numbers).

The exponential of $X \,\epsilon\, \mathcal{L}$ is denoted as $\mathbf{exp}(X)$ and given by

$$\mathbf{exp}(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!} \tag{65}$$

If **L** be a matrix Lie group then $X$ must have and inverse and $\det(X) \neq 0$, so there exists a matrix $\alpha$, logarithm of $X$ s.t.

$$X = \mathbf{exp}(\alpha) = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \tag{66}$$

The set of all matrices $\alpha$ whose exponentials belong to a group **L** is known as the Lie algebra, $\mathcal{L}$ of **L**. Using the definition of the exponential series, we have

$$X = \lim_{k \to \infty} \left( I + \frac{\alpha}{k} \right)^k \text{ and } \alpha = \lim_{k \to \infty} k(X^{1/k} - I) \tag{67}$$

For large k, the matrix (I + $\alpha$/k) is an operator of the group and gives $X$ by iteration. It is an **infinitesimal operator**, since for large k it differs from the identity operator by an infinitesimal amount. Lie group is a manifold with a group structure and Lie algebra is its **tangent space**. Elements of Lie algebra to a Lie group are sometimes referred to as **generators** of the group. The Lie algebra can be thought of as the **infinitesimal vectors/operators** generating the group (at least locally) by means of exponential map, but the Lie algebra doesn't form a generating set in strict sense.

### 4.8.4 Generators of Lie Group

Let $\alpha_j$ be a set of parameters the parameterize the Lie group **L** such that

$$g(\alpha)|_{\alpha=0} = e \tag{68}$$

where e is the identity element of Lie group **L**.

Let D be a representation of **L** then the linear operators of the representation will be parameterized the same way,

$$D(\alpha)|_{\alpha=0} = I \tag{69}$$

where I is the identity operator.

Then in some neighborhood of the identity element, we can Taylor expand D($\alpha$) as

$$D(d\alpha) = 1 + id\alpha_a X_a + .... \tag{70}$$

where d$\alpha$ is infinitesimal and the sum over repeated indices should be understood as the "Einstein summation convention".

Then the generators of **L** are

$$X_a = -i\frac{\partial D}{\partial \alpha_a}(\alpha)\Big|_{\alpha=0} \tag{71}$$

In such a case,

$$D(\alpha) = e^{iX_a\alpha^a} \tag{72}$$

Why do we want to study Lie group through their generators? The truth is that the generators form an Algebra under **matrix commutation**. First of all, it is observed that the generators must all be operators in the same vector space the representation of the Lie Group maps to, so they must also form an operator vector space just like the group elements themselves. The only thing we need to check is that there exists a bilinear product in the vector space of the generators. The fact is that there is such a bilinear product called matrix commutator.

Let two group elements e$^{i\alpha^a X_a}$ and e$^{i\beta^b X_b}$, where $\alpha^a$ and $\beta^b$ are just two different set of parameters for the Lie group, then by group closure,

$$e^{i\alpha^a X_a}e^{i\beta^b X_b} = e^{i\delta^c X_c} \tag{73}$$

Then we can write,

$$i\delta^c X_c = ln(e^{i\alpha^a X_a}e^{i\beta^b X_b}) = ln(1 + e^{i\alpha^a X_a}e^{i\beta^b X_b} - 1) = ln(1 + K) \tag{74}$$

where $K = e^{i\alpha^a X_a}e^{i\beta^b X_b} - 1$

Then performing binomial expansion we finally get,

$$i\delta^c X_c = i\alpha^a X_a + i\beta^b X_b - \alpha^a X_a \beta^b X_b - \frac{1}{2}\Big((\alpha^a X_a)^2 - (\beta^b X_b)^2\Big) + \frac{1}{2}\Big(\alpha^a X_a + \beta^b X_b\Big)^2 + .... \tag{75}$$

where the term

$$\Big((\alpha^a X_a)^2 - (\beta^b X_b)^2\Big) + \Big(\alpha^a X_a + \beta^b X_b\Big)^2 = [\alpha^a X_a, \beta^b X_b] \tag{76}$$

known as commutation operation. No matter what the infinite sum in the right-hand side will be, it must be some linear combination of the generators multiply by $i$. Lets call it the linear combination $\gamma^c X_c$. Then,

$$\begin{aligned} [\alpha^a X_a, \beta^b X_b] &= i\gamma^c X_c = i\alpha^a \beta^b f^c_{ab} X_c \\ \implies [X_a, X_b] &= i f^c_{ab} X_c \end{aligned} \tag{77}$$

The commutation relation 77 is called the Lie algebra of the group which are completely determined by the constants $f^c_{ab}$ called the structural constants of the Lie Algebra. Even if we call them constants, $f^c_{ab}$ is really a three-tensor of constants.

We might wonder if we keep expanding 74 beyond second order, we would need additional conditions to make sure that the group multiplication law is maintained. The remarkable thing is that we don't. The commutator relation 77 is enough. In fact with structural constants $f^c_{ab}$ we can reconstruct as accurately as we like for any $\alpha$ and $\beta$ in some finite neighborhood of the origin. Thus the $f^c_{ab}$ are very important-they summarize virtually the entire group multiplication law.

For arbitrarily given $X_a$ , $X_b$ , $X_c$ , we can map from any two of the three generators to another one using matrix commutation. Thus, matrix commutator forms a bilinear product of the vector space of generators. The matrix commutators '[ , ]' are often called **Lie Bracket** in order to distinguish with another possible bilinear product for the generators **Poisson Bracket** '{ , }'. In physics, Poisson Bracket is often used when we are working with classical Hamilton's equations. When the Poisson Bracket is replaced by Lie Bracket, it automatically implies the Hamiltonian or the Lagrangian is mapped to some Hilbert Space where the matrix algebra works and the full Hamiltonian equation of motion reduces to the **Heisenberg picture** of quantum mechanics.

### 4.8.5 SU(2) and SU(3) Groups

Unit rotation in 2-dimension is the SO(2) group given by

$$R(\theta) = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \tag{78}$$

However, a 3-dimensional representation of SO(2) is different from SO(3). An SO(2) group in $\mathbb{R}^3$ is

$$R(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & -sin\theta \\ 0 & sin\theta & cos\theta \end{pmatrix} \tag{79}$$

which is a unit rotation along the axis of unit vector (1, 0, 0) and this group doesn't represent akk 3D unit rotations. There doesn't exist a 2D representation of a 3D rotation but there exists a 2D representation of some group whose algebra is the same as SO(3), the group is called SU(2), which is defined in complex space.

To represent SO(2) using complex number we first define

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{80}$$

then equation 78 can be represented as

$$R(\theta) = cos\theta + sin\theta \tag{81}$$

This equation forms a Lie group called U(1) where real number '1' is the generator of U(1) since

$$e^{i*1\theta} = cos\theta + isin\theta \tag{82}$$

We can easily build a complex version of the SO(2) algebra because the entire symmetry group relies only on two degrees of freedom, so we can effectively use a single complex number to capture the dof. However, two dof is not enough for SO(3), so we consider higher dimension which leads us to U(2). We can claim that only those unitary operators with determinant 1 will be sufficient to model the Algebra of SO(3) given that SO(3) has determinant 1 as well. The unitary operators with determinant 1 are called **Special Unitary** operators and they form the group SU(2).

$$SU(2) = \left\{ \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha \end{pmatrix} : \alpha, \beta \, \epsilon \, \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\} \tag{83}$$

The implication of the matrix elements of SU(2) in the usual two dimensional representation is that SU (2) gives a maximum of four degrees of freedom, with two degree of freedom for each complex number. We then need four basis operators to write all operators in SU(2).

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; i = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \; j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; k = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{84}$$

where $i^2 = j^2 = k^2 = -1$. and every other element $m$ of SU(2) can be written as

$$m = a + bi + cj + dk \tag{85}$$

this is called **quaternion**. We can then express our SO(3) basis operators $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ as

$$\begin{aligned} t_x(\theta) &= cos(\theta) + sin(\theta)i \\ t_y(\theta) &= cos(\theta) + sin(\theta)j \\ t_z(\theta) &= cos(\theta) + sin(\theta)k \end{aligned} \tag{86}$$

SU(2) is a dual cover of SO(3) which means we can map exactly two elements of SU(2) i.e. two $t$'s in SU(2) to one $R(\theta)$ in SO(3), so if we consider a convex combination of $t_x$ , $t_y$ , and $t_z$, we span all elements of the entire SO(3). We can examine the generators of the above 2D representation of SU(2):

$$\begin{aligned} \sigma_x &= -i\frac{dt_x}{d\theta}(0) = -i \begin{pmatrix} -sin(0) & icos(0) \\ icos(0) & -sin(0) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_y &= -i\frac{dt_y}{d\theta}(0) = -i \begin{pmatrix} -sin(0) & cos(0) \\ -cos(0) & -sin(0) \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_z &= -i\frac{dt_z}{d\theta}(0) = -i \begin{pmatrix} -sin(0) + icos(0) & 0 \\ 0 & -sin(0) - icos(0) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{87}$$

These are called **Pauli matrices** and for $i \neq j \neq k$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{88}$$

Thus, the structural constants of SU(2) and SO(3) algebra must be the same.

For example, the spin 1/2 representation is

$$\begin{aligned} J_1^{1/2} &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\sigma_1}{2} \\ J_2^{1/2} &= \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\sigma_2}{2} \\ J_3^{1/2} &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\sigma_3}{2} \end{aligned} \tag{89}$$

The spin 1/2 representation is the simplest representation of SU(2). It is called the **defining representation** of SU(2), and is responsible for the name Special Unitary (SU).

# 5 PARTICLE PHYSICS

## 5.1 Discrete Symmetries

**J$^{\mathbf{PC}}$**

## 5.2 Spatial Rotations

**Clebsch-Gordan coefficients** generator of rotations Wigner D-matrix

Helicity conservation in $e^+e^- \rightarrow \mu^+\mu^-$

Helicity, $\sigma. \frac{p}{|p|}$, is the projection of the spin along the momentum direction.

## 5.3 Lorentz Invariance

Most high-energy physics requires energy scales E >> m$_p$, so it is essential that the requirements of special relativity be respected. In practice, this means identifying suitable 4-vectors and Lorentz invariants. Although the position and direction of particles is important when actually performing experiments, the results are most often derived from knowledge of the energy and momentum of the interacting particles.

## 5.4 Rapidity

High-energy hadron-hadron interactions tend to produce final states with limited transverse momentum with respect to the initial beam direction. In such circumstances, rapidity (y) and transverse mass (m$_T$) are convenient variables.

$$y = \frac{1}{2} ln\Big(\frac{E + p_z}{E - p_z}\Big), \qquad m_T = \sqrt{m^2 + p_T^2} \tag{90}$$