Group Theory

Mapping : Let S_1 and S_2 be two sets. A mapping (or **function**) $f : S_1 \to S_2$ is defined if :

$$\forall s_1 \in S_1, \ \exists!\ s_2 = f(s_1) \in S_2.$$

Note that, we did **not** claim that : $\forall s_2 \in S_2, \ \exists\ s_1 \in S_1$, such that $s_2 = f(s_1)$.
If however that is true, we call 'f' an **onto** or a **surjective** map.

We have said, for every s_1 in S_1, there should be one and **only one** s_2 in S_2. If you find more that one element in S_2, such that $s_2 = f(s_1)$, then by definition, 'f' is not a mapping at all. Sometimes we loosely talk about a multi-valued function. However a multi-valued function is no function.

Ex : $y = \sqrt{x}$ is not a function from $\mathbb{R} \to \mathbb{R}$ since, for a –ve x, there is no y. It is not even a function from $\mathbb{R}_+ \to \mathbb{R}$, since, for x = 25, we have f(x) = 5 and – 5, both $\in \mathbb{R}$.
However, $y = +\sqrt{x}$, or, $-\sqrt{x}$ is a function alright, as you have given the instruction of accepting only one sign.
Alternatively, $y = \sqrt{x}$ is not a function from $\mathbb{R}_+ \to \mathbb{R}_+$, as the –ve roots $\notin \mathbb{R}_+$/

Although we have demanded that for every s_1, there should be only one s_2, we did **not** demand that two distinct elements of S_1 cannot be mapped to a single s_2.
If however that is the case, we call the map 'f' a **one-one** or **injective** map.

If a map which is both surjective and injective, it is called a **bijective map**.

A **Cartesian product** of two sets $S_1 \times S_2$ is defined as a set of **all** ordered duplets of the form : (s_1, s_2) where, $s_1 \in S_1$, $s_2 \in S2$.

A **binary operation**, defined on a set S is nothing but a map from $S \times S \to S$
i.e., corresponding to every pair of elements $s_1, s_2 \in S$, we associate an element $s \in S$ and we denote it as : $s = s_1 \circ s_2$.

A '**groupoid**' is non-empty set G, with a binary operation '$\circ$' defined on it.

$$\forall g_1, g_2 \in G, \ \exists!\ g \in G \text{ such that } g = (g_1 \circ g_2) \qquad (1)$$

The fact that $\forall g_1, g_2 \in G$, $(g_1 \circ g_2) \in G$ is known as the '**closure property**'. Note that the way we have defined the binary operation, the closure property is built in.

. A '**semi-group**' is a **groupoid** G, with the binary operation '$\circ$' obeying the '**associative rule'**, i.e.,

$$\forall g_1, g_2, g_3 \in G, \ (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3) \qquad (2)$$

A '**monoid**' is a **semi-group** G, with an '**identity element**' i, such that :

$$\forall g \in G, \ (i \circ g) = (g \circ i) = g. \qquad (3)$$

A '**group**' is a **monoid** G, such that every element of G has an '**inverse**' g^{-1} obeying :

$$\forall g \in G, \ (g^{-1} \circ g) = (g \circ g^{-1}) = i \qquad (4)$$

Actually, if $(i \circ g) = g$, i is called the left identity and if $(g \circ i) = g$, i is called the right identity. However, for a group, the left identity and the right identity are identical. Similarly, if $(g^{-1} \circ g) = i$, g^{-1} is called the left inverse and if $(g \circ g^{-1}) = i$, g^{-1} is called the right inverse, but again if G is a group, the left inverses and the right inverses are identical.

If, in addition to the rules (1) to (4), the binary operation '∘' obeys the '**commutative rule**' :

$$\forall g_1, g_2 \in G, \quad (g_1 \circ g_2) = (g_2 \circ g_1), \qquad (5)$$

then G is called a **commutative** or an **Abelian group**. We emphasize that the commutative property is **not essential** for a group.

Examples :

The set of all natural numbers $\mathbb{N}$ {1, 2, 3, · · ·} forms a semi-group under common addition.

The set of all natural numbers, together with 0 {0, 1, 2, 3, · · ·} forms a monoid under addition.

The set of all integers $\mathbb{Z}$ {1, 2, 3, · · ·, 0, – 1, – 2 · · ·} forms a abelian group under addition.

The set of all natural numbers $\mathbb{N}$ forms a monoid under common multiplication.

The set of all integers $\mathbb{Z}$ also forms a monoid under multiplication.

The set of all rational numbers $\mathbb{Q}$ (which can be expressed as the ratio of two integers p/q) is a monoid under multiplication, but $\mathbb{Q} - \{0\}$ forms abelian group under multiplication.

$\mathbb{R} - \{0\}$ and $\mathbb{C} - \{0\}$ (where $\mathbb{R}$ and $\mathbb{C}$ are the set of all real and complex numbers respectively) also form abelian groups respectively under real and complex multiplication.

The set of all three dimensional vectors form an abelian group under vector addition.

The set of all $m \times n$ matrices form an abelian group under matrix addition.

The set of all $n \times n$ matrices forms a monoid under matrix multiplication. If you have the set of all $n \times n$ *invertible* matrices, i.e., with non-zero determinant, then a non-abelian group is formed.

Some elementary Properties

1) The left cancellation of property : $g \circ a = g \circ b \Rightarrow a = b$

   Proof :

   $$g \circ a = g \circ b$$

   Multiply both sides by g^{-1} from left $\Rightarrow g^{-1} \circ (g \circ a) = g^{-1} \circ (g \circ b)$

   $$\Rightarrow (g^{-1} \circ g) \circ a = (g^{-1} \circ g) \circ b$$

   $$\Rightarrow i \circ a = i \circ b$$

   $$\Rightarrow a = b$$

   Similarly, we can establish the right cancellation property.

2) The identity is unique : if $(i_1 \circ g) = g$ and $(i_2 \circ g) = g$, then $i_1 = i_2$.

   Proof :

   $$(i_1 \circ g) = g \text{ and } (i_2 \circ g) = g \Rightarrow i_1 \circ g = i_2 \circ g$$

   Hence, by right cancellation of g : $i_1 = i_2$

3) The inverse is unique : if $(g_1^{-1} \circ g) = i$ and $(g_2^{-1} \circ g) = i$, then $g_1 = g_2$.

   Proof :

   $$(g_1^{-1} \circ g) = i \text{ and } (g_2^{-1} \circ g) = i \Rightarrow (g_1^{-1} \circ g) = (g_2^{-1} \circ g)$$

   Hence, by right cancellation of g : $g_1^{-1} = g_2^{-1}$

4) $(g^{-1})^{-1} = g$

   Proof :

   $$g^{-1} \circ g = i \text{ as } g^{-1} \text{ is the inverse of } g$$

$$g^{-1} \circ (g^{-1})^{-1} = i \text{ as } (g^{-1})^{-1} \text{ is the inverse of } (g^{-1})$$
$$\Rightarrow g^{-1} \circ g = g^{-1} \circ (g^{-1})^{-1}$$

Hence, by left cancellation of g^{-1} : $g = (g^{-1})^{-1}$

Group Multiplication Table

The result of the binary operation between all the pairs of elements can be tabulated in the following way :

$\circ$	g_1	g_2	g_3
g_1	$g_1 \circ g_1$	$g_1 \circ g_2$	$g_1 \circ g_3$
g_2	$g_2 \circ g_1$	$g_2 \circ g_2$	$g_2 \circ g_3$
g_3	$g_3 \circ g_1$	$g_3 \circ g_2$	$g_3 \circ g_3$

Here we have followed the convention that the elements from the leftmost column is placed on the left and that from the topmost row is placed on the right.

Re-arrangement theorem :

Each element of the group will appear once and only once in each row and each column of the table.

Reason :

Suppose the element g_n doesn't appear in the first row. That means there is no element 'g' such that : $g_1 \circ g = g_n$.

Choose : $g = g_1^{-1} \circ g_n \Rightarrow g_1 \circ (g_1^{-1} \circ g_n) = (g_1 \circ g_1^{-1}) \circ g_n = g_n$.

Next, suppose that the element g_n is missing in the first column, i.e., we should look for a 'g' such that : $g \circ g_1 = g_n$, but that is ready at hand!

Choose : $g = g_n \circ g_1^{-1} \Rightarrow (g_n \circ g_1^{-1}) \circ g_1 = g_n$.

Now suppose that an element g_n appears twice in the first row. That means, there are two distinct elements g and g′, such that :

$g_1 \circ g = g_n$ and $g_1 \circ g' = g_n$

but then $g_1 \circ g = g_1 \circ g' \Rightarrow g = g'$ [by left cancellation of g_1]

Similar logic applies for the columns.

Application :

The above theorem helps filling up the group multiplication table and hence figure out the group structure through a crossword puzzle game. Let a group have only three elements : *i*, *a*, *b*. If we reserve the first row and the first column for the identity operations, the group multiplication table looks like :

$\circ$	*i*	*a*	*b*
i	*i*	*a*	*b*

a	a	?	?
b	b	?	?

What is ? a and b are already present in that column and that row. So, '?' is i.
What is ? b and i are already present in that row. So, '?' is a.
What is ? a and i are already present in that column. So, '?' is b.
What is ? a and b are already present in that column and that row. So, '?' is i.

Thus the multiplication table reads :

$\circ$	i	a	b
i	i	a	b
a	a	b	i
b	b	i	a

From the second row, one may say that $b = a^2$ and $a^3 = i$. Here of course a^2 means $a \circ a$, a^3 means $(a \circ a) \circ a$, and so on.

Cyclic Group:

If the elements of a group can be expressed as the powers of one of the elements, then the group is called a cyclic group and the element is referred as the 'generator' of the group. We have an example immediately above. The elements of the group are : i, a, a^2, while $a^3 = i$.

The generator of a cyclic group need not be unique. For example, in the above group, both a and a^2 can act as the generator.

Remember a chapter called 'Indices' in your school-days' Algebra ? The 'laws of indices' you learned, mostly holds for a cyclic group, e.g.,

$a^3 \circ a^2 = (a^3) \circ (a \circ a) = (a^3 \circ a) \circ a = a^4 \circ a = a^5$ → addition of powers

$(a^3)^2 = (a^3) \circ (a \circ a \circ a) = (a^3 \circ a) \circ (a \circ a) = (a^4 \circ a) \circ a = a^6$

→ multiplication of powers

$a^3 \circ a^{-2} = (a \circ a \circ a) \circ (a^{-1} \circ a^{-1}) = (a \circ a) \circ (a \circ a^{-1}) \circ a^{-1} = (a \circ a) \circ a^{-1}$

$= a \circ (a \circ a^{-1}) = a$ → subtraction of powers, etc.

Homo-morphism and Iso-morphism :

Let $(G_1, \circ)$ and $(G_2, *)$ be two groups, with respect to the binary operations '$\circ$' and '$*$'. A mapping $f : G_1 \to G_2$ is called a homo-morphism if :

$$\forall g_1, g_2 \in G_1, \quad f(g_1 \circ g_2) = f(g_1) * f(g_2)$$

Note that : $f(g_1)$ and $f(g_2) \in G_2$, so the operation between them has to be '$*$'.
One says that the '**binary operation and the mapping commutes**' i.e., you can perform any of them first. (Of course the binary operation changes before and after the mapping.)

Example :

Consider the cube roots of unity : $1, \omega, \omega^2$. Remember that they can be expressed as :

$1,\ e^{2\pi i/3},\ e^{4\pi i/3}$, since, $e^{2\pi i} = 1$. These three numbers form a group under complex multiplication. Consider next the three matrices which represent rotations through
$0,\ 2\pi/3$ and $4\pi/3$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} \cos 120^\circ & \sin 120^\circ \\ -\sin 120^\circ & \cos 120^\circ \end{bmatrix} \qquad \begin{bmatrix} \cos 240^\circ & \sin 240^\circ \\ -\sin 240^\circ & \cos 240^\circ \end{bmatrix}$$

These matrices form a group under matrix multiplication. It is easy to check that the two groups are homo-morphic.
Consider a group $(G_2, *)$ with only one element i, the only possible operation being :
$i * i = i$. Then for any other group $(G_1, \circ)$, the mapping : $\forall g \in G_1,\ \ f(g) = i$
is a homo-morphism. You can check this easily.
If the homo-morphism map is onto and one-one, i.e., bijective, then the homo-morphism is called iso-morphism. The example given immediately above is a homo-morphism but **not** an iso - morphism. The example given before that is a homo-morphism as well as an iso-morphism.
It is said that '**in an iso-morphism the group structure is preserved**', i.e., the group multiplication table remains the same.

Problems :

1) Prove that if $f : G_1 \to G_2$ is a homo-morphism, then : $f(i_1) = i_2$, where i_1 and i_2 are the identity elements of the two groups.
2) Prove that if $f : G_1 \to G_2$ is a homo-morphism, then : $\forall g \in G_1\ \ f(g^{-1}) = \{f(g)\}^{-1}$.

Subgroup

If $(G, \circ)$ is a group an H is a subset of G, then $(H, \circ)$ is called a subgroup if $(H, \circ)$ forms a group itself, i.e.,

$H \subseteq G$ and $\forall h_1, h_2 \in H,\ \exists!\ h \in H$ such that $h = (h_1 \circ h_2)$: 'closure property'
$\forall h_1, h_2, h_3 \in H,\ (h_1 \circ h_2) \circ h_3 = h_1 \circ (h_2 \circ h_3)$: 'associative property'*
the identity element $i \in H$: existence of identity
$\forall h \in H,\ h^{-1} \in H$: existence of inverse.

*the associative property here however is obvious, because it is followed by all elements of G. Hence, the elements of H follow it trivially.

Every group G has two trivial subgroups. One is the group itself and another is a single-element group $\{i\}$. G itself is of course a subset of G (called an **improper subset**) and obviously, satisfies all the requirements of a group. $\{i\}$ also is a subset and it is easy to check that the above four properties are all obeyed.

Theorem : the following is a **necessary and sufficient** condition for $H \subseteq G$ to be a subgroup of G.

$$\forall h_1, h_2 \in H,\ h_1 \circ h_2^{-1} \in H.$$

Proof of necessity : This is easy !

$$h_2 \in H \ \Rightarrow\ h_2^{-1} \in H \ \Rightarrow\ h_1 \circ h_2^{-1} \in H.$$

Proof of sufficiency :

1) Take $h_2 = h_1$. $h_1 \circ h_1^{-1} \in H \Rightarrow i \in H$.
2) Take $h_1 = i$, then $\forall h \in H$, $h_1 \circ h^{-1} \in H \Rightarrow i \circ h^{-1} \in H \Rightarrow h^{-1} \in H$.
3) For any $h_1, h \in H$, take $h_2 = h^{-1} \Rightarrow h_1 \circ (h^{-1})^{-1} \in H \Rightarrow h_1 \circ h \in H$.

Note that the order of the steps is important here, e.g., you cannot prove (2) before (1).

Matrix Representation of Groups

Let $(\mathbb{G}, \circ)$ be any group with respect to the binary operation '$\circ$' and $\mathbb{M}$ be a group of $n \times n$ square matrices with respect to the standard matrix multiplication. Then a homomorphism $f : \mathbb{G} \to \mathbb{M}$ is called an **n-dimensional matrix representation** of $\mathbb{G}$.

$$\forall g_1, g_2 \in \mathbb{G}, \text{ let } f(g_1) \text{ be } M_1 \text{ and } f(g_2) \text{ be } M_2. \text{ Then } f(g_1 \circ g_2) = M_1 M2.$$

This maps every element of the abstract group $\mathbb{G}$ to a matrix and replaces the abstract operation '$\circ$' by matrix multiplication, bringing us to our familiar world.
Although not mentioned, it is clear that the matrices must all be **non-singular** (with a non-zero determinant), because $\forall g \in \mathbb{G}$, $f(g) = M \Rightarrow f(g^{-1}) = M^{-1}$, which doesn't exist if M is singular.

If the mapping 'f' is bijective, i.e., an isomorphism, then the representation is called '**faithful**'. Under an unfaithful representation, more than one group element may be mapped to a single matrix. For example, if all the elements of a group (any group) is mapped to the $n \times n$ unit matrix, we get an n-dimensional representation of the group. In a faithful representation however, each group element goes to one and only one matrix $\in \mathbb{M}$.

Example :

Consider a group with four elements : i, a, b, ab with the properties : $a^2 = b^2 = (ab)^2 = i$. This is known as **Klien's four-group**. You may think of the reflection about the x-axis as 'a', the reflection about the y-axis as 'b' and the a rotation by 180^o about the z-axis as (ab). A little thought will reveal that these are the possible operations on a rectangle that doesn't produce any visible change. Now map these elements to following matrices :

$$i \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; a \to \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \; b \to \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \; ab = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

It is easy to check that we have found out a two-dimensional representation of the four-group.
Is the representation faithful ? Yes, it is.

Let $f : \mathbb{G} \to \mathbb{M}$ be an n-dimensional matrix representation of $\mathbb{G}$ and S be any $n \times n$ non-singular matrix, not necessarily belonging to $\mathbb{M}$. Let for any $g \in \mathbb{G}$, $f(g) = M$.
We define : $N = S^{-1}MS$. We know that the transformation : $M \to S^{-1}MS$ is called a similarity transformation. The collection of all such matrices :

$$\mathbb{N} = \{N \mid N = S^{-1} f(g) S\}$$

forms another matrix group and the mapping

$$f' : \mathbb{G} \to \mathbb{N} \text{ such that } N = S^{-1} f(g) S$$

forms another homomorphism and hence another representation of $\mathbb{G}$. This is called an **equivalent representation**.

To show that f ′ is a homomorphism :

$$\forall g_1, g_2 \in \mathbb{G},\ N_1 = S^{-1} f(g_1) S = S^{-1} M_1 S \text{ and } N_2 = S^{-1} f(g_2) S = S^{-1} M_2 S$$

$$\begin{aligned} \Rightarrow\ N_1 N_2 &= S^{-1} M_1 S S^{-1} M_2 S \\ &= S^{-1} M_1 M_2 S \\ &= S^{-1} f(g_1 \circ g_2) S \\ &= f'(g_1 \circ g_2). \end{aligned}$$

Reducible Representation :

Let a set of matrices $\mathbb{M}(\mathbb{G})$ form a n-dimensional representation of the group $\mathbb{G}$. If you can find an equivalent representation $\mathbb{N}(\mathbb{G})$ with the help of a matrix S, such that the members of $\mathbb{N}(\mathbb{G})$ reduces to the partitioned form :

$$N(g)_{n\times 5} = \begin{bmatrix} A(g)_{m\times m} & 0_{m\times p} \\ B(g)_{p\times m} & D(g)_{p\times p} \end{bmatrix}, \text{ where } n = m + p,$$

then $\mathbb{M}(\mathbb{G})$ is said to be a **reducible representation** of $\mathbb{G}$. If no such equivalent representation can be found then the representation $\mathbb{M}(\mathbb{G})$ is **irreducible**. One can show that in the former case, the matrices A(g) and D(g) themselves form an $m \times m$ and a $p \times p$ representation of $\mathbb{G}$ respectively.

Proof :

$$\text{Let } N(g_1)_{n\times n} = \begin{bmatrix} A(g_1)_{m\times m} & 0_{m\times p} \\ B(g_1)_{p\times m} & D(g_1)_{p\times p} \end{bmatrix} \text{ and } N(g_2)_{n\times n} = \begin{bmatrix} A(g_2)_{m\times m} & 0_{m\times p} \\ B(g_2)_{p\times m} & D(g_2)_{p\times p} \end{bmatrix}$$

According to the rules of multiplication of partitioned matrices :

$$N(g_1)\, N(g_2) = \begin{bmatrix} A(g_1)A(g_2) & 0 \\ B(g_1)A(g_2) + D(g_1)B(g_2) & D(g_1)D(g_2) \end{bmatrix}$$

Clearly, the product $N(g_1)\, N(g_2)$ has the required form. Since $N(\mathbb{G})$ is a representation,

$N(g_1)\, N(g_2) = N(g_1 \circ g_2)$, which, by our assumption, must be of the form :

$$N(g_1 \circ g_2) = \begin{bmatrix} A(g_1 \circ g_2) & 0 \\ B(g_1 \circ g_2) & D(g_1 \circ g_2) \end{bmatrix}$$

If we focus at the sub-matrices A(g) and D(g), we find that : $A(g_1 \circ g_2) = A(g_1)A(g_2)$ and $D(g_1 \circ g_2) = D(g_1)D(g_2)$, which shows that they form two lower dimensional representations themselves.

Consider the example of a two dimensional representation of the 4-group, given above. Note that each matrix is of the form we have just mentioned with $B(g) = 0\ \forall g$, Thus the above representation is reducible. It can be split up into two one dimensional representations $A(\mathbb{G})$ and D(G) as :

$A(i) = 1,\ A(a) = -1,\ A(b) = 1,\ A(ab) = -1$ and

$D(i) = 1,\ D(a) = 1,\ A(b) = -1,\ A(ab) = -1.$

However, these representations are clearly not faithful.

Representation Space

We have learnt, that to construct a representation, one has to find a matrix corresponding to very element of a group. One might wonder how should one look for such a collection of matrices ? A natural way of finding them is to find a representation space first.

A **group action** is defined on a set S, if every element 'g' of a group (G, $\circ$) induces a map g : S $\rightarrow$ S such that $\forall s \in S$:

(i) $g_1(g_2(s)) = (g_1 \circ g_2)(s)$ and
(ii) $i(s) = s$, where i is the identity element of G.

Next, consider a vector space $\mathbb{V}(\mathbb{F})$ in place of a general set S, where G acts on $\mathbb{V}(\mathbb{F})$.
If in addition to the above two properties, we further have

(iii) $\forall v \in \mathbb{V}$ and $\forall g \in G : g(v_1 + v_2) = g(v_1) + g(v_2)$ and
(iv) $\forall v \in \mathbb{V}$, $\forall \lambda \in \mathbb{F}$ and $\forall g \in G : \lambda\, g(v) = g(\lambda v)$,

then $\mathbb{V}(\mathbb{F})$ is a **representation space** for G, which 'carries' a representation of G.
Consider a basis set $\{e_1, e_2, e_3, \ldots\}$ for $\mathbb{V}$. For a particular group element 'g',
$g(e_i)$ will be an element of $\mathbb{V}$, which can be expanded again, in terms of the given basis as : $g(e_i) = \Sigma_j\, e_j\, T_{ji}(g)$, where $T_{ji}(g)$ are the expansion coefficients.
The map : $g \rightarrow T_{ji}(g)$ forms a representation for g.
Proof :

Let $g_1(e_i) = \Sigma_j\, e_j\, T_{ji}$ and $g_2(e_i) = \Sigma_j\, e_j\, S_{ji}$,
then, $(g_1 \circ g_2)(e_i) = g_1(g_2(e_i))$ [by (i)]

$= g_1\left(\Sigma_j\, e_j\, S_{ji}\right)$

$= \Sigma_j\, g_1(e_j)\, S_{ji}$ [by (iii) and (iv)]

$= \Sigma_j\, \Sigma_k\, e_k\, T_{kj}\, S_{ji}$

$= \Sigma_k\, e_k\, (TS)_{ki}$ [by matrix multiplication rule]

Thus $(g_1 \circ g_2)$ is mapped to the matrix TS, demonstrating the homo-morphism.
In general, a vector $v \in \mathbb{V}$ can be expanded as : $v = \Sigma_i\, e_i\, x_i$, where x_i are the expansion coefficients. Under the action of the group element 'g' this is mapped to :

$$g(v) = g\left(\Sigma_i\, e_i\, x_i\right) = \Sigma_i\, g(e_i)\, x_i = \Sigma_i\, \Sigma_j\, e_j\, T_{ji}(g)\, x_i$$

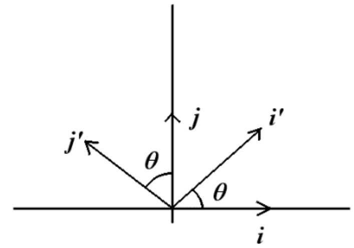
The coefficients x_i are thus mapped to : $\Sigma_i\, T_{ji}(g)\, x_i$
Clearly, the dimension of the vector space becomes the dimension of the representation.
Example :

Consider the x-y plane as $\mathbb{V}$. The set $\{\boldsymbol{i}, \boldsymbol{j}\}$ forms basis in this space.
Consider all rotations about the z-axis as our group G. A particular rotation by an angle 'θ' takes $\boldsymbol{i}$ to $\boldsymbol{i}'$, which can be expanded in terms of $\{\boldsymbol{i}, \boldsymbol{j}\}$ as :

$$\boldsymbol{i}' = \cos\theta\, \boldsymbol{i} + \sin\theta\, \boldsymbol{j} = \boldsymbol{i}\, T_{11} + \boldsymbol{j}\, T_{21}$$

Similarly under the rotation, $\boldsymbol{j}$ goes to $\boldsymbol{j'}$, which can be expanded as :

$$\boldsymbol{j'} = -\sin\theta\, \boldsymbol{i} + \cos\theta\, \boldsymbol{j} = \boldsymbol{i}\, T_{12} + \boldsymbol{j}\, T_{22}$$



Thus the matrix representation of the rotation becomes :

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

<u>Change of Basis and Equivalent Representation :</u>

Suppose now we make a change of basis :

$e_i \rightarrow e_i' = \Sigma_j\, e_j\, R_{ji}$, where 'R' is the transformation matrix.

$$g(e_i') = g(\Sigma_j\, e_j\, R_{ji}) = \Sigma_j\, g(e_j)\, R_{ji} \quad \text{[by (iii) and (iv)]}$$

$$= \Sigma_j\, (\Sigma_k\, e_k\, T_{kj})\, R_{ji}\,, \text{ where } g(e_j) = \Sigma_k\, e_k\, T_{kj}$$

$$= \Sigma_k\, e_k\, (TR)_{ki}$$

Now this can again be expanded in terms of the new basis $\{e_i'\}$.

$$e_i' = \Sigma_k\, e_k\, R_{ki} \Rightarrow e_k = \Sigma_r\, e_r'\, (R^{-1})_{rk}$$

$$\Rightarrow g(e_i') = \Sigma_k\, e_k\, (TR)_{ki} = \Sigma_k\, \Sigma_r\, e_r'\, (R^{-1})_{rk}\, (TR)_{ki}$$

$$= \Sigma_r\, e_r'\, (R^{-1}TR)_{ri}$$

Thus, in place of 'T', g is represented by another matrix $R^{-1}TR$, which is an equivalent representation.

<u>Invariant Subspace and Reducible Representation :</u>

Let $\mathbb{V}(\mathbb{F})$ carry an n-dimensional representation of G. If we find an m-dimensional (m < n) sub-space $\mathbb{W}(\mathbb{F})$ of $\mathbb{V}(\mathbb{F})$ such that

$$\forall v \in \mathbb{W} \text{ and } \forall g \in G : g(v) \in \mathbb{W}$$

then $\mathbb{W}$ is called an **invariant subspace** of $\mathbb{V}$ and the representation carried by $\mathbb{V}$ is reducible.

<u>Proof :</u>

Start with a basis $\{f_i \mid i = 1, 2, \ldots m\}$ of $\mathbb{W}$ and append it to form a basis for $\mathbb{V}$ as $\{ f_1, f_2, f_3, \ldots f_m, e_1, e_2, e_3, \ldots e_p\}$, where m + p = n.

For any $g \in G$, $g(f_i)$ will be of the form : $\Sigma_j\, f_j\, A_{ji}$, i, j = 1, 2, . . . m

For example, $g(f_1)$ may be written as :

$$f_1\, A_{11} + f_2\, A_{21} + f_3\, A_{31} + \ldots + e_1\, 0 + e_2\, 0 + e_3\, 0 + \ldots$$

e_k's on the other hand will be mapped to :

$$g(e_k) = \Sigma_i\, f_i\, B_{ik} + \Sigma_l\, e_l\, D_{lk}\,, \quad i = 1, 2, \ldots m, \quad k, l = 1, 2, \ldots p$$

For example, $g(e_1)$ may be written as :

$$f_1\, B_{11} + f_2\, B_{21} + f_3\, B_{31} + \ldots + e_1\, D_{11} + e_2\, D_{21} + e_3\, D_{31} + \ldots$$

Thus, the matrix representation of the element 'g' will be of the form :

$$\lceil A(g)_{m\times m} \quad 0_{m\times p} \rceil \,, \text{ where } n = m + p,$$

$\lfloor B(g)_{p\times m} \quad D(g)_{p\times p} \rfloor$

As a concrete example, let's suppose $\mathbb{V}$ is a5-dimensional vector space and $\mathbb{W}$ is a 3-dimensional invariant subspace of $\mathbb{V}$. A basis set of W is { f_1, f_2, f_3}, which has been extended to {f_1, f_2, f_3, e_1, e_2} to form a basis of $\mathbb{V}$.
Since $g(f_1) \in \mathbb{W}$, it can be expanded as : $f_1 A_{11} + f_2 A_{21} + f_3 A_{31}$,
which can be written as :

$g(f_1) = f_1 A_{11} + f_2 A_{21} + f_3 A_{31} + e_1 C_{11} + e_2 C_{21}$, where $C_{11} = C_{21} = 0$

similarly,

$g(f_2) = f_1 A_{12} + f_2 A_{22} + f_3 A_{32} + e_1 C_{12} + e_2 C_{22}$, where $C_{12} = C_{22} = 0$

and

$g(f_3) = f_1 A_{13} + f_2 A_{23} + f_3 A_{33} + e_1 C_{13} + e_2 C_{23}$, where $C_{13} = C_{23} = 0$

However, $g(e_1)$ and $g(e_2) \notin \mathbb{W}$, so, in their expansion, all the coefficients of e_1 and e_2 won't vanish. These expansions will look like :

$g(e_1) = f_1 B_{11} + f_2 B_{21} + f_3 B_{31} + e_1 D_{11} + e_2 D_{21}$

and

$g(e_2) = f_1 B_{12} + f_2 B_{22} + f_3 B_{32} + e_1 D_{12} + e_2 D_{22}$

Now, one can easily read the read out the elements of the matrix A, B, C and D.

<u>The Great Orthogonality Theorem</u>

Now we are going to prove a very powerful theorem on irreducible representations. It states that if $T_\alpha(G)$ and $T_\beta(G)$ are two **irreducible representations** of a finite group G, with dimensions 'm' and 'n', then :

$$\Sigma_{g\in G}\ T_\alpha(g)_{ij}\, T_\beta(g^{-1})_{kl} = (|g| / n)\ \delta_{\alpha\beta}\, \delta_{il}\, \delta_{jk} ,$$

where $|g|$ stands for the order (no. of elements) of G. Here $\alpha \neq \beta$ signifies that the two representations are inequivalent.
The proof follows from the two Schur's Lemma, that we discussed the other day. Let us remind ourselves what the two lemmas taught us.

<u>Schur's Lemma 1 :</u>

In an irreducible representation T(G) of a finite group G, if all the matrices T(g) ($g \in G$) commute with a matrix P, then P is a multiple of the identity matrix.

<u>Schur's Lemma 2 :</u>

If $T_1(G)$ and $T_2(G)$ are two irreducible representations of a finite group G and M is a matrix such that $\forall g\in G$: $T_1(g)\, M = M\, T_2(g)$, then either M is the null matrix, or, M is invertible.

We shall first apply Lemma 2. Let us construct the matrix M as :

$$M = \Sigma_{g\in G}\ T_1(g)\, A\, T_2(g^{-1}),$$

where $T_1(G)$ and $T_2(G)$ are two inequivalent representations and 'A' is an arbitrary matrix, so far. However, if $T_1(g)$ are $m \times m$ matrices and $T_2(g)$ are $n \times n$ matrices, then 'A' has to be of the order $m \times n$.

Multiply both sides by $T_1(g_1)$ from left and $T_2(g_1^{-1})$ from right (where g_1 is just another element of G) :

$$\begin{aligned} T_1(g_1)\ M\ T_2(g_1^{-1}) &= T_1(g_1)\ [\Sigma_{g\in G}\ T_1(g)\, A\, T_2(g^{-1})]\ T_2(g_1^{-1}) \\ &= \Sigma_{g\in G}\ T_1(g_1)\, T_1(g)\, A\, T_2(g^{-1})\, T_2(g_1^{-1}) \\ &= \Sigma_{g\in G}\ T_1(g_1 g)\, A\, T_2(g^{-1} g_1^{-1}) \\ &= \Sigma_{g\in G}\ T_1(g_1 g)\, A\, T_2[(g_1 g)^{-1}] \\ &= \Sigma_{g'\in G}\ T_1(g')\, A\, T_2[(g')^{-1}], \text{ where } g_1 g = g' \end{aligned}$$

(The summation over g' and that over g are just re-arrangement of the terms.)

Thus, $T_1(g_1)\, M\, T_2(g_1^{-1}) = M \Rightarrow \mathbf{T_1(g_1)\, M = M\, T_2(g_1)}$

Hence, by Schur's Lemma-2, $\mathbf{M = 0}$, or, $\Sigma_{g\in G}\ T_1(g)\, A\, T_2(g^{-1}) = 0$

$$\text{i.e., } \Sigma_{g\in G}\ T_1(g)_{ij}\, A_{jk}\, T_2(g^{-1})_{kl} = 0$$

So far, A was arbitrary. Now let us choose A such that a specific element of A (say A_{pq}) equals 1 and the others are zero, i.e., $A_{jk} = \delta_{jp}\,\delta_{kq}$

$$\Rightarrow M_{il} = \Sigma_{g\in G}\ T_1(g)_{ij}\, \delta_{jp}\, \delta_{kq}\, T_2(g^{-1})_{kl} = 0$$

$$\Rightarrow \boldsymbol{\Sigma_{g\in G}\ T_1(g)_{ip}\, T_2(g^{-1})_{ql} = 0\ \ \forall\, i, p, q, l.} \text{ ----- (1)}$$

For example, if only $A_{25} = 1$ and all other elements of $A = 0$, then :

$\Sigma_{g\in G}\ T_1(g)_{i2}\, T_2(g^{-1})_{5l} = 0\ \ \forall$ i and l.

Next, when $T_1(G)$ and $T_2(G)$ are the same representation, say $T(G)$, we apply Schur's Lemma-1.

Now, $\mathbf{T(g_1)\, P = P\, T(g_1) \Rightarrow P = \lambda}\, I$,

$$\text{or, } \Sigma_{g\in G}\ T(g)\, A\, T(g^{-1}) = \lambda\, I \text{ ----- (2)}$$

$$\text{i.e., } \Sigma_{g\in G}\ T(g)_{ij}\, A_{jk}\, T(g^{-1})_{kl} = \lambda\, \delta_{il}$$

If we choose $A_{jk} = \delta_{jp}\,\delta_{kq}$ again. then :

$$\Sigma_{g\in G}\ T(g)_{ij}\, \delta_{jp}\, \delta_{kq}\, T(g^{-1})_{kl} = \lambda\, \delta_{il}$$

$$\Rightarrow \Sigma_{g\in G}\ T(g)_{ip}\, T(g^{-1})_{ql} = \lambda\, \delta_{il} \text{ ----- (3)}$$

Taking the trace of both sides of (2) :

$$\text{Tr}\ \Sigma_{g\in G}\ T(g)\, A\, T(g^{-1}) = n\,\lambda, \text{ where 'n' is the dimension of } T(g)$$

$$\Rightarrow \text{Tr}\ \Sigma_{g\in G}\ T(g^{-1})\, T(g)\, A = n\,\lambda, \text{ by the cyclic property of Trace}$$

$$\Rightarrow \text{Tr}\ \Sigma_{g\in G}\ T(g^{-1}g)\, A = n\,\lambda,$$

but $(g^{-1}g) = i$, the identity element of the group and $T(i) = I$, the identity matrix.

$$\Rightarrow \Sigma_{g\in G}\ \text{Tr}\, A = |g|\ \text{Tr}\, A = n\,\lambda$$

Substituting back in (3) : $\Sigma_{g\in G}\ T(g)_{ip}\, T(g^{-1})_{ql} = (|g|/n)\,(\text{Tr}\, A)\,\delta_{il}$

We have chosen the (p, q)-th element of A = 1 and other elements = 0.

Obviously, if $p \neq q$, (Tr A) = 0 and if $p = q$, (Tr A) = 1

$$\Rightarrow \boldsymbol{\Sigma_{g\in G}\ T(g)_{ip}\, T(g^{-1})_{ql} \ =\ (|g|/n)\, \delta_{pq}\, \delta_{il}}$$

Combining with (1), when the two representations were inequivalent, one may say that :

$$\boldsymbol{\Sigma_{g\in G}\ T_\alpha(g)_{ip}\, T_\beta(g^{-1})_{ql} \ =\ (|g|/n)\, \delta_{\alpha\beta}\, \delta_{pq}\, \delta_{il}}$$

If we consider unitary representations, as one may do without much loss of generality,

$$\Sigma_{g\in G}\ T_\alpha(g)_{ip}\, T_\beta(g^{-1})_{ql} \ =\ \Sigma_{g\in G}\ T_\alpha(g)_{ip}\, [T_\beta(g)^{-1}]_{ql} \ ==\ \Sigma_{g\in G}\ T_\alpha(g)_{ip}\, [T_\beta(g)^{\dagger}]_{ql}$$

$$\boldsymbol{=\ \Sigma_{g\in G}\ T_\alpha(g)_{ip}\, T_\beta{}^*(g)_{lq} \ =\ (|g|/n)\, \delta_{\alpha\beta}\, \delta_{il}\, \delta_{pq}}$$

Orthogonality of Characters :

Set $i = p$ and sum over that index and then set $l = q$ and sum over that index too.

At first, we have : $\Sigma_{g\in G}\ T_\alpha(g)_{pp}\, T_\beta{}^*(g)_{ql} \ =\ (|g|/n)\, \delta_{\alpha\beta}\, \delta_{pl}\, \delta_{pq} \ =\ (|g|/n)\, \delta_{\alpha\beta}\, \delta_{lq}$

and then : $\Sigma_{g\in G}\ T_\alpha(g)_{pp}\, T_\beta{}^*(g)_{qq} \ =\ (|g|/n)\, \delta_{\alpha\beta}\, \delta_{qq} \ =\ |g|\ \delta_{\alpha\beta}$ [since $\delta_{qq} = n$]

Essentially, we have taken the Traces of the matrices $T_\alpha(g)$ and $T_\beta(g)$. The traces of the matrices of a representation are called the '**characters**'. Denoting them by $\chi(g)$, we have the Orthogonality Theorem for characters :

$$\boldsymbol{\Sigma_{g\in G}\ \chi_\alpha{}^{(g)}\, \chi_\beta{}^{(g)*} \ =\ |g|\, \delta_{\alpha\beta}}$$

Lie Group and Lie Algebra

An **infinite group** is one which has an infinite number of elements, e.g., the set of all integers (forming a group under arithmetic addition), the set of all rotation matrices in a plane (forming a group under matrix multiplication). This number of elements may be a countable or an uncountable infinity. In the first example, the number is a countable infinity (where the element of a set may be put in an one-one correspondence with the set of Natural Numbers), in the second example, it is uncountable.

Continuous group : We shall assume that our group elements may be parametrized by a set of real numbers, i.e., there is a one-one correspondence between the group elements (g) and the set of parameters (x_1, x_2, …). The number of parameters is neither insufficient (to uniquely specify a group element) nor excess. The set of parameters (say n) belong to the n-dimensional, real inner product space. We call it the **parameter space**. If some of these parameters (at least one) vary continuously over a range, we say that the group is a **continuous group**. Note that the term 'continuous' is used here in the colloquial sense, not referring to any 'continuous mapping' defined in calculus.

Topological group : Usually, it is difficult to develop a calculus and discuss limit and continuity over an abstract group, because we shall require the concept of 'nearness', which is difficult to introduce for an abstract group. What would you mean, for example, by saying that two matrices or two Quantum Mechanical operators are 'close' to each other? That is why we talk in terms of the n-dimensional, real parameter space, which has an in-built topology (concept of nearness). Now, there are two standard operations defined over a group, multiplying (we refer to the binary operation of the group) one element by another and taking inverse of an element.

Each of these define a map over the group :

(i) For every $g \in (G, *)$, $g_1 * g = g_2$ defines a map : $g_1 \rightarrow g_2$

[You could also define a map using the left multiplication : $g * g_1 = g_2$]

(ii) and another map is $g \rightarrow g^{-1}$

The group is called a **Topological group, if the maps (i) and (ii) are continuous**. However, to take advantage of the topology of the parameter space, we shall say that the maps $\mathbf{r}(g_1) \rightarrow \mathbf{r}(g_2)$ and $\mathbf{r}(g) \rightarrow \mathbf{r}(g^{-1})$ are continuous, where $\mathbf{r}(g)$ is the vector in the parameter space which corresponds to the group element g.

For (i) : $\forall\ \varepsilon > 0$ but however small, $\exists\ \delta > 0$ such that $|\,\mathbf{r}(g_1) - \mathbf{r}(g_1')\,| < \delta \Rightarrow |\,\mathbf{r}(g_2) - \mathbf{r}(g_2')\,| < \varepsilon$

For (ii) : $\forall\ \varepsilon > 0$ but however small, $\exists\ \delta > 0$ such that $|\,\mathbf{r}(g) - \mathbf{r}(g')\,| < \delta \Rightarrow |\,\mathbf{r}(g^{-1}) - \mathbf{r}\{(g')^{-1}\}\,| < \varepsilon$

Connectedness : If the parameter space is connected, we say the group is **connected**, which means any two points in the parameter space may be joined by **a continuous path which lies entirely within the parameter space**. Note that if some of the parameters of the group take discrete values, the group will have disconnected components.

If two points in the parameters can be connected by more than one set of paths, such that a path of one set **cannot be continuously deformed** into any path of another set, the parameter space (and the group) is called **multiply connected**. In technical language, such continuous deformation is called **Homotopy**.

Compactness : A group is said to be compact, if its parameter space is compact. We avoid the general Topological definition of compactness. For our purpose, the parameter space is compact if it is **closed and bounded**. Many result for finite groups hold for these compact groups, e.g.,

a) for every representation, we can find an equivalent unitary representation;

b) the unitary reducible representations are completely reducible, i.e., they can be brought in a block diagonal form;

c) all irreducible representations are finite dimensional.

Lie group : In a Topological group, if we can find a neighbourhood of the identity element, within which **the maps (i) and (ii) mentioned above** (the group multiplication and taking inverse) **are differentiable, the group is called a Lie group** (named after the Norwegian mathematician Sophus Lie).

Generators : A generator is basically the derivatives of a group element, with respect to its parameters. A group is usually parametrized in a way, such that the identity element corresponds to the origin of the parameter space, i.e., parametrized as $(0, 0, 0, \cdots)$. Let us consider a group of matrices or a matrix representation of an abstract group, with elements like $g(x_1, x_2, \cdots)$.

then we define :

$$\tau_i = \partial g/\partial x_i = \lim_{\delta x_i \rightarrow 0} [g(0, 0, \delta x_i, 0, \cdots) - g(0, 0, 0, 0, \cdots)] / \delta x_i$$

$$= \lim_{\delta x_i \rightarrow 0} [g(0, 0, \delta x_i, 0, \cdots) - I] / \delta x_i$$ as the generators of the group.

Thus, for small δx_i's : $g(\delta x_1, \delta x_2, \cdots) = I + \Sigma\, \tau_i\, \delta x_i$

If the parameters x_i's are not small, we can take $\delta x_i = x_i/N$, where δx_i's are small and group element $g(x_1, x_2, \cdots)$ may be considered as a product of N group elements of the form : $g(\delta x_1, \delta x_2, \cdots)$.

For example, a rotation through a finite angle θ, may be considered as N successive rotations, each with an angle θ/N.

Thus, for finite parameters : $g(x_1, x_2, \cdots) = [I + \Sigma \tau_i (x_i/N)]^N$

We can make (x_i/N) as small as we like by making 'N' large and in the limit $N \to \infty$:

$$g(x_1, x_2, \cdots) = \exp[\Sigma\tau_i x_i] = \exp[\tau_1 x_1 + \tau_2 x_2 + \tau_3 x_3 + \cdots]$$

If we are working with a unitary representation, 'g' is unitary $\Rightarrow \tau_i$' s are anti-Hermitian.

Instead, we may prefer to define τ_i' s as :

$$\boldsymbol{\tau_i = \lim_{\delta x_i \to 0} [g(0, 0, \delta x_i, 0, \cdots) - g(0, 0, 0, 0, \cdots)] / i\, \delta x_i}$$

$\Rightarrow$ for small δx_i's : $g(\delta x_1, \delta x_2, \cdots) = I + i \Sigma \tau_i \delta x_i$

and for finite parameters : $\boldsymbol{g(x_1, x_2, \cdots) = \exp[i \Sigma \tau_i x_i]}$. **---- (1)**

If 'g' is unitary now, τ_i' s will be Hermitian.

Example : The rotation matrices about the z-axis form a group. The matrices are of the form :

$$g(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \approx \begin{bmatrix} 1 & \theta \\ -\theta & 1 \end{bmatrix}, \text{ for small } \theta.$$

They may be parametrized by the rotation angle 'θ', where $g(\theta = 0)$ is the identity matrix.

So, $g(\theta) - g(0) \approx \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix} \Rightarrow \lim_{\theta \to 0} [g(\theta) - g(0)] / i\theta = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, which is nothing but the Pauli matrix σ_2.

In eq.(1), the expression $[\Sigma \tau_i x_i]$ in the power forms an n dimensional, real vector space '$\mathbb{L}$' (for n parameters), with τ_i's acting as the bases.

Algebra : A vector space $\mathbb{V}(\mathbb{F})$ is called an **Algebra**, if in addition to the vector addition and the multiplication with scalars, we have another binary operation $\mathbb{V} \times \mathbb{V} \to \mathbb{V}$, which we denote as : $a \times b$ ($a, b \in \mathbb{V}$), with the properties :

(1) **Closure** : $\forall a, b \in \mathbb{V}, a \times b \in \mathbb{V}$

(2) **Linearity** : $\forall a, b, c \in \mathbb{V}$ and $\lambda \in \mathbb{F}$, $a \times (\lambda b + \mu c) = \lambda (a \times b) + \mu (a \times c)$

also, $(\lambda a + \mu b) \times c = \lambda (a \times c) + \mu (b \times c)$.

If, in addition to the above two rules, the **commutative** rule is also obeyed, i.e.,

(3) $a \times b = b \times a$, $\forall a, b \in \mathbb{V}$, then $\mathbb{V}$ is called an 'commutative algebra' and on obeying the **associative** rule :

(4) $(a \times b) \times c = a \times (b \times c)$ $\forall a, b, c \in A$, A becomes a 'associative algebra'.

Example : The Complex Algebra is a commutative, associative Algebra. The Matrix Algebra, under the standard matrix multiplication, is an associative, but non-commutative Algebra.

Lie Algebra : Instead, a vector space $\mathbb{V}(\mathbb{F})$ is called an **Lie Algebra,** if we have a binary operation $[a, b]$ defined on it (which we call the Lie multiplication) with the properties :

(1) **Closure** : $\forall a, b \in \mathbb{V}, [a, b] \in \mathbb{V}$

(2) **Anti-commutativity** : $[a, b] = -[b, a]$

(3) **Linearity** : $\forall a, b, c \in \mathbb{V}$ and $\lambda \in \mathbb{F}$, $[a, [\lambda b + \mu c]] = \lambda [a, b] + \mu [a, c]$

(4) **Jacobi Identity** : $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$.

Our familiar vector cross product is an example of Lie multiplication.

If $g(x, 0, 0, \cdots) = \exp(\tau_1 x)$ and $g(0, y, 0, \cdots) = \exp(\tau_2 y)$, then their product must belong to the group. Now, by **Baker-Hausdorff theorem** :

$e^A e^B = e^C$, where $C = A + B + ½ [A, B] + 1/12 [A, [A, B]] + 1/12 [B, [A, B]] + \cdots$

So, $\exp(\tau_1 x) \exp(\tau_2 y) = e^C$, where $C = \tau_1 x + \tau_2 y + xy/2 [\tau_1 , \tau_2] + \cdots$,

where the higher order terms may be neglected for small x and y. Now, $\tau_1 x$ and $\tau_2 y$ already $\in$ **L**,

so, $[\tau_1 , \tau_2]$ must also $\in$ **L**.

Thus the closure property is ensured. The criteria (2) to (4) are properties of for matrix commutators in general, So, if we consider the commutator bracket as the Lie multiplication, the vector space **L** mentioned above, becomes a Lie Algebra. Your familiar vector cross product is also another example of Lie multiplication.

By the closure property of **L**, $[\tau_a, \tau_b] \in$ **L** and hence, can be expressed in terms of the bases :

$$[\tau_a, \tau_b] = i \Sigma_c f_{abc} \tau_c \quad \text{---- (2)}$$

where the f_{abc}'s are called the '**structure constants**'. The factor 'i' has been introduced for making the f_{abc}'s real. Note that if, the generators are Hermitian, their commutator is anti-Hermitian and so is $(i\tau_c)$ on the RHS of (2).

Clearly f_{abc}'s are antisymmetric in a, b (look at the LHS). We shall show, that under suitable condition, they are completely antisymmetric in all the three indices.

Example : The Pauli matrices provide a certain representation for the generators of the group SU(2) (as we shall show later).

$[\sigma_a/2, \sigma_b/2] = i\, \varepsilon_{abc}\, \sigma_c/2 \;\Rightarrow$ the structure constants are the Levi-civita symbols.

Taking the complex conjugation of both sides of (2), we find that :

$[\tau_a^*, \tau_b^*] = -i f_{abc} \tau_c^* \;\Rightarrow\; [(-\tau_a^*) , (-\tau_b^*)] = i f_{abc} (-\tau_c^*)$

Thus if we can find a matrix representation for the generators τ_a's for a group, then the matrices $(-\tau_a^*)$ will provide another representation for the generators and hence, for the group. If this is an equivalent representation, i.e., if we can find a matrix s, such that :

$$\mathbf{s\, \tau_a\, s^{-1} = (-\tau_a^*)}, \text{ **we call it a 'real repersentation'**}$$

We know that $(\sigma_2)^{-1} = \sigma_2$. So, Taking σ_2 as 's' in case of the above representation for SU(2), it is easy to check that :

$$\sigma_2\, \sigma_1\, \sigma_2^{-1} = \sigma_2\, \sigma_1\, \sigma_2 = -\,\sigma_1 = -\,\sigma_1^*,$$
$$\sigma_2\, \sigma_2\, \sigma_2^{-1} = \sigma_2\, \sigma_2\, \sigma_2 = \sigma_2 = -\,\sigma_2^*,$$
$$\sigma_2\, \sigma_3\, \sigma_2^{-1} = \sigma_2\, \sigma_3\, \sigma_2 = -\,\sigma_3 = -\,\sigma_3^*,$$

Showing that this representation is real.

From eqn.(2) :

$$[\tau_d, [\tau_a , \tau_b]] = [\tau_d, i f_{abc} \tau_c] = i f_{abc} [\tau_d, \tau_c] = -\, f_{abc} f_{dce} \tau_e$$

Hence, by Jacobi identity :

$[\tau_d, [\tau_a , \tau_b]] + [\tau_a, [\tau_b , \tau_d]] + [\tau_b, [\tau_d , \tau_a]] = f_{abc} f_{dce} \tau_e + f_{bdc} f_{ace} \tau_e + f_{dac} f_{bce} \tau_e = 0,$

[Just replace $a \to b \to d \to a$, cyclically]

$$\text{or, } \mathbf{(f_{abc} f_{dce} + f_{bdc} f_{ace} + f_{dac} f_{bce}) = 0} \quad \text{---- (3)}$$

$$\Rightarrow f_{abc} f_{dce} - f_{dbc} f_{ace} + f_{adc} f_{cbe} = 0 \quad \text{[by antisymmetry]}$$

$$\text{or, } f_{abc} f_{dce} - f_{dbc} f_{ace} = -f_{adc} f_{cbe} \quad \text{---- (4)}$$

If we define a set of matrices T_a, such that $\mathbf{(T_a)_{bc} = -if_{abc}}$, then (4) $\Rightarrow$

$$(iT_a)_{bc} (iT_d)_{ce} - (iT_d)_{bc} (iT_a)_{ce} = -f_{adc} (iT_c)_{be}$$

$$\text{or, } (T_aT_d)_{be} - (T_dT_a)_{be} = if_{adc} (T_c)_{be} \Rightarrow [T_a , T_d] = i f_{adc} T_c$$

showing that $(T_a)_{bc} = -if_{abc}$ provides a representation for the generators. This is known as the '**adjoint representation**'.

Clearly, $\mathbf{(-T_a^*)_{bc} = (T_a)_{bc}} \Rightarrow$ **the adjoint representation is real.**

SU(2) :

The set of all n × n unitary matrices form a group. The group is called **U(n)**. The set of 'special' unitary matrices form a sub-group of U(n), where the adjective 'special' signifies that their determinant is unity. The sub-group is named **SU(n)**. Similarly, the group formed by all n × n orthogonal matrices is named **O(n)** and those with determinant = 1 among them, are together called **SO(n)**. You can check for yourself, that U(n), SU(n), O(n) and SO(n) all form groups.

A unitary matrix U may be written as e^{iH}, where H is Hermitian. So, if U is a 2 × 2 unitary matrix, then H is a 2 × 2 Hermitian matrix. If you were tempted, let's warn that the set of all n × n Hermitian matrices **do not** form a group. H has the general form :

$$H = \begin{bmatrix} z_1 & x - iy \\ x + iy & z_2 \end{bmatrix},$$ where x, y, z_1, z_2 are real parameters. Instead of z_1 and z_2 as

Parameters, we may choose to work with $\xi = (z_1 + z_2)/2$ and $\eta = (z_1 - z_2)/2$, so that z_1 becomes $(\xi + \eta)$ and z_2 becomes $(\xi - \eta)$ an H looks like :

$$H = \begin{bmatrix} \xi + \eta & x - iy \\ x + iy & \xi - \eta \end{bmatrix} = \xi \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \eta \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$= \xi \mathbf{I} + \eta \boldsymbol{\sigma_z} + x \boldsymbol{\sigma_x} + y \boldsymbol{\sigma_y}$$

and U assumes the form : $\exp \{i (\xi \mathbf{I} + \eta \boldsymbol{\sigma_z} + x \boldsymbol{\sigma_x} + y \boldsymbol{\sigma_y})\}$. Clearly, the set of all 2 × 2 Hermitian matrices form a four-parameter Lie algebra, I, σ_z, σ_x and σ_y being the generators.

Now, $\mathbf{\det (U) = \det (e^{iH}) = e^{Tr (iH)}}$.

You can show this easily, working in a basis, where H is diagonal, since then :

$$iH = \text{diag } (i\lambda_1, i\lambda_2, \cdots) \Rightarrow \text{Tr } (iH) = (i\lambda_1 + i\lambda_2 + \cdots)$$

$$\Rightarrow e^{\text{Tr } (iH)} = \exp (i\lambda_1) \exp (i\lambda_2) \cdots = \det [\text{diag } \{\exp (i\lambda_1), \exp (i\lambda_2), \cdots\}] = \det U.$$

Hence, det (U) = 1 $\Rightarrow$ Tr (H) = 0. Therefore, if U $\in$ SU(2), H is of the form :

$$H = \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix} = z \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$= z \boldsymbol{\sigma_z} + x \boldsymbol{\sigma_x} + y \boldsymbol{\sigma_y},$$

which shows that the Pauli matrices act as the three generators of SU(2). The above representation of the SU(2) algebra is called the **'fundamental representation'**.

We know that : $[\sigma_i , \sigma_j] = 2i \varepsilon_{ijk} \sigma_k \Rightarrow \mathbf{[\sigma_i /2 , \sigma_j /2] = i \varepsilon_{ijk} \sigma_k /2.}$

So, if $(\sigma_i /2)$'s are chosen as generators, then the **Levi-civita symbols are the structure constants**.

The above commutation relation is familiar to us. It appears in angular momentum algebra, $[L_i , L_j] = i \varepsilon_{ijk} L_k$. We also derived that $L^2 = (L_1^2 + L_2^2 + L_3^2)$ commutes with all the L_i's.

If a matrix (operator) commutes with all the generators of a Lie algebra, it is called a '**Casimir operator**'(named after the Dutch Physicist Hendrik Casimir).

In our case, $(\sigma_1/2)^2 + (\sigma_2/2)^2 + (\sigma_3/2)^2 = ¾\ I$, which plays the role of L^2.
We note that : ¾ = ½ (½ + 1), which means that $(\sigma_i/2)$ behave as the angular momenta of a particle with $\ell = ½$.

Guided by our old experience, we can introduce ladder operators (matrices) :
$\boldsymbol{\sigma_\pm = (\sigma_1/2 \pm i\, \sigma_2/2)}$ and obtain the raising/lowering effect :

$$\sigma_+|\ell, m\rangle = \{(\ell - m)(\ell + m + 1)\}^{½}|\ell, m+1\rangle,\ \sigma_-|\ell, m\rangle = \{(\ell + m)(\ell - m + 1)\}^{½}|\ell, m-1\rangle, \quad \text{---- (v)}$$

with $\ell = ½$ and $m = \pm ½$.

Direct Sum : If we find two representations of a group, we can construct a representation by taking a direct sum of the matrices. The resulting matrices will obviously be in a block diagonal form and the representation space will be (m + n) dimensional, if the dimensions of original representations were 'm' and 'n'. If M(g) and N(g) are the matrices representing a group element 'g' in the two representations, then the corresponding matrix in the direct sum representation will be :

$$M\oplus N(g) = \begin{bmatrix} M(g)_{m\times m} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & N(g)_{n\times n} \end{bmatrix}$$

Direct, or, Tensor Product : We can also construct a new representation by the taking the direct product of the matrices, which we call the **product representation**. The representation space will be (m × n) dimensional. The matrix in the product representation $P(g) = M\otimes N(g)$ is defined as : $P(g)_{i\alpha,\, j\beta} = M(g)_{ij}\, N(g)_{\alpha\beta}$. Note that if i, j have 'm' possible values, say 1, 2, 3 and α, β have 'n' possible values, say 1, 2, then $i\alpha$, or, $j\beta$ will have (m × n) possible values like : 11,12, 21, 22, 31, 32.

If the elements of the representation space for M(g) are denoted as $|i\rangle, |j\rangle$, etc., and those of the representation space for M(g) is denoted as $|\alpha\rangle, |\beta\rangle$, etc., then elements of the representation space for the product representation look like : $|i\rangle|\alpha\rangle$, $|j\rangle|\beta\rangle$, etc. The way P(g) acts on, say $|i\rangle|\alpha\rangle$ is :

$$\mathbf{P(g)|i\rangle|\alpha\rangle = M(g)|i\rangle\, N(g)|\alpha\rangle.}$$

If the generators of the first representation are denoted by $X^{(a)}$, and those of the second representation by $Y^{(a)}$, then the generators of the product representation are given by :

$$\boldsymbol{[Z^{(a)}]_{i\alpha,\, j\beta} = [X^{(a)}]_{ij}\, \delta_{\alpha\beta} + \delta_{ij}\, [Y^{(a)}]_{\alpha\beta}},$$

So that, while acting on a vector $|i\rangle|\alpha\rangle$, the first term attacks the ket $|i\rangle$, leaving $|\alpha\rangle$ unaffected and the second term attacks the ket $|\alpha\rangle$, leaving $|i\rangle$ unaffected. Sometimes we loosely write the above equation as : $\boldsymbol{Z^{(a)} = X^{(a)} + Y^{(a)}}$.

A Product Representation of SU(2) : Let us find the product representation of two fundamental representations of SU(2). This is basically finding the angular momenta of a system of two spin - ½ particles. The basis vectors of the representation space for the product representation are of the form : $|\ell_1, m_1\rangle|\ell_2, m_2\rangle$. Following the above notation, the generators of the product representation L_i may be written in terms of those of the original representations $L_i^{(1)}$ and $L_i^{(2)}$ as :

$J_i = L_i^{(1)} + L_i^{(2)}$, where $\{L_x^{(1)}\}_{ij} = (\sigma_x)_{ij}$ etc., and $\{L_x^{(2)}\}_{\alpha\beta} = (\sigma_x)_{\alpha\beta}$ etc.

A new set of basis of the representation space for the product representation is chosen as the eigen vectors of J^2 and J_z, which are of the form : $|j, m_j\rangle$.

The eigen vector with the highest eigen value of J_z is $|j, 1\rangle = |½, ½\rangle|½, ½\rangle$.

[Since, $J_z = L_z^{(1)} + L_z^{(2)}$, $L_z^{(1)}|½, ½\rangle|½, ½\rangle$ equals ½ $|½, ½\rangle|½, ½\rangle$

$L_z^{(2)}|½, ½\rangle|½, ½\rangle$ also equals ½ $|½, ½\rangle|½, ½\rangle$

$\Rightarrow\ J_z|½, ½\rangle|½, ½\rangle = 1\ |½, ½\rangle|½, ½\rangle$.]

What is the value of 'j' ?

Note that : $\mathbf{J}^2 = \{\mathbf{L}^{(1)}\}^2 + \{\mathbf{L}^{(2)}\}^2 + 2\,\mathbf{L}^{(1)}\cdot\mathbf{L}^{(2)}$

$= \{\mathbf{L}^{(1)}\}^2 + \{\mathbf{L}^{(2)}\}^2 + 2\{L_x^{(1)} L_x^{(2)} + L_y^{(1)} L_y^{(2)} + L_z^{(1)} L_z^{(2)}\}$

$= \{\mathbf{L}^{(1)}\}^2 + \{\mathbf{L}^{(2)}\}^2 + \{L^{(1)}_+ L^{(2)}_- + L^{(1)}_- L^{(2)}_+\} + 2\,L_z^{(1)} L_z^{(2)}$

$\Rightarrow \mathbf{J}^2|½, ½\rangle|½, ½\rangle = (¾ + ¾ + 0 + 0 + 2 \times ¼)\ |½, ½\rangle|½, ½\rangle$

$= 2\ |½, ½\rangle|½, ½\rangle = 1\,(1 + 1)|½, ½\rangle|½, ½\rangle$.

Thus j = 1. We can get the other eigenvectors with the same 'j' but different 'm' by lowering :

$J_-\ |1, 1\rangle = (L^{(1)}_- + L^{(2)}_-)\ |½, ½\rangle|½, ½\rangle$

$\Rightarrow\ \sqrt{2}|1, 0\rangle = \{\,|½, -½\rangle|½, ½\rangle + |½, ½\rangle|½, -½\rangle\}$

[$\because J_-|j, m\rangle = \{(j + m)(j - m + 1)\}^{½}|j, m-1\rangle$]

$\Rightarrow\ \mathbf{|1, 0\rangle = \{|½, -½\rangle|½, ½\rangle + |½, ½\rangle|½, -½\rangle\}/\sqrt{2}}$

Applying J_- again, : $\sqrt{2}|1, -1\rangle = \{\,|½, -½\rangle|½, -½\rangle + |½, -½\rangle|½, -½\rangle\}/\sqrt{2}$

[$\because L^{(1)}_-\ |½, -½\rangle|½, ½\rangle = 0$ and $L^{(2)}_-\ |½, ½\rangle|½, -½\rangle = 0$]

$\Rightarrow\ \mathbf{|1, -1\rangle = |½, -½\rangle|½, -½\rangle}$

$|1, 1\rangle$, $|1, 0\rangle$ and $|1, -1\rangle$ will span a 3-dimensional subspace, which will furnish a 3-dimensional representation of SU(2). However, m = 0 is also possible for j = 0. The vector $|0, 0\rangle$ will be orthogonal to $|1, 0\rangle$, since they are eigenvectors corresponding to different eigenvalues of J^2. One can find a vector perpendicular to $|1, 0\rangle$ simply by inspection :

$\mathbf{|1, 0\rangle = \{|½, -½\rangle|½, ½\rangle - |½, ½\rangle|½, -½\rangle\}/\sqrt{2}}$.

The one dimensional subspace spanned by this vector will furnish a one dimensional representation. Thus we have split our 4-dimensional representation space into a direct sum of a 3-dim and a 1-dim representation spaces. Symbolically, one expresses it as : **2⊗2 = 3⊕1**.