LIE GROUPS, LIE ALGEBRAS, AND APPLICATIONS IN PHYSICS

JOO HEON YOO

ABSTRACT. This paper introduces basic concepts from representation theory, Lie group, Lie algebra, and topology and their applications in physics, particularly, in particle physics. The main focus will be on matrix Lie groups, especially the special unitary groups and the special orthogonal groups. They play crucial roles in particle physics in modeling the symmetries of the subatomic particles. Of the many physical applications, the paper will introduce two concepts/phenomena known as 'Isospin' and 'Eightfold Way'.

CONTENTS

1. PRELIMINARIES

Here, we define basic concepts that will be used later on. Note that the scalar field of the vector space will be the complex number, $\mathbb{C}$, unless mentioned otherwise.

Definitions 1.1. Given a vector space V over field $\mathbb{C}$, a **norm** on V is a function $\rho : V \to \mathbb{R}$ with the following properties:
For all $a \in \mathbb{C}$ and all $x, y \in V$,

1) $\rho(ax) = |a|\rho(x)$,

2) $\rho(x + y) \leq \rho(x) + \rho(y)$,

3) $\rho(x) = 0 \implies x = 0$.

The particular norm, which we will use in this paper, for a vector $x \in \mathbb{C}^n$ will be denoted by $\|x\|$ and defined as following:

$$\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Date: September 17, 2015.

Whereas the norm of an $n \times n$ complex matrix, A, will be defined as following:

$$\|A\| = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|}.$$

It is not hard to check that these two norms satisfy the three properties mentioned above.

Definitions 1.2. Let V be a vector space with basis, $\{v_i\}_{i=1}^n$. Then, the *symmetric product*, $\mathrm{Sym}^2 V$, is defined as

$$\mathrm{Sym}^2 V = V \otimes V / (v_i \otimes v_j \sim v_j \otimes v_i),$$

with $1 \le i, j \le n$. The *alternating product*, $\bigwedge^2 V$, is defined as

$$\bigwedge^2 V = V \otimes V / (v_i \otimes v_j \sim -v_j \otimes v_i),$$

with $1 \le i, j \le n$.

Definition 1.3. A **group** is a set G together with an operation, $\cdot : G \times G \to G$, satisfying four requirements known as the *group axioms*:

1) $\forall\ a, b \in G,\ \ a \cdot b \in G,$

2) $\forall\ a, b, c \in G,\ \ a \cdot (b \cdot c) = (a \cdot b) \cdot c,$

3) $\exists\ e \in G$ such that $a \cdot e = e \cdot a = a\ \forall\ a\ \in G,$

4) for each $a \in G,\ \exists\ a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e.$

2. Representation Theory

Here, we define representation and its associated concepts and state some examples. Before starting, it could be helpful to first understand why the concept of representation is important. A representation can be thought of as an action of group on a vector space. Such actions can arise naturally in mathematics and physics, and it is important to study and understand them. For example, consider a differential equation in three-dimensional space with a rotational symmetry. If the equation has a rotational symmetry, then the solution space will be invariant under rotations. Thus, the solution space will constitute of a representation of the rotation group $SO(3)$. Hence, knowing what all the representations of $SO(3)$ are, it is easy to narrow down what the solution space of the equation will be. In fact, one of the main applications of representations is exploiting the system's symmetry. In a system with a symmetry, the set of symmetries form a group, and the representation of this symmetry group allows you to use that symmetry to simplify the given system. We will see more of these applications, particularly in physics, in the final section.

Definition 2.1. A **representation** of a group G on a vector space V is defined as a homomorphism $\phi : G \to GL(V)$. To each $g \in G$, the representation map assigns a linear map, $\rho_g : V \to V$. Although V is actually the representation space, one may, for short, refer to V as the representation of G.

Definition 2.2. A **subrepresentation** of a representation V is a vector subspace W of V, which is invariant under G. This means $\mathrm{Im}(W) = W$ under the action of each $g \in G$.

Definition 2.3. A representation V is called **irreducible** if there is no proper nonzero invariant subspace W of V. That is, it has no subrepresentation, except itself and the trivial space.

Now that the basic concepts are defined, we can look at some examples of representation.

Example 2.4. (*Trivial Representation*) Every element $g \in G$ gets mapped to the identity mapping between the vector space V and itself. Hence, all elements of G acts as the identity on all $v \in V$.

Example 2.5. (*Standard Representation*) If we let G be S_n, the symmetric group on n elements, then G is obviously represented by vector space $V \cong \mathbb{C}^n$ with n basis vectors: each element in S_n permutes the basis vectors. Now, there is, one dimensional subrepresentation W of V spanned by sum of the basis vectors. The **standard representation** is $n-1$ irreducible representation V/W, the quotient space.

Example 2.6. (Dual Representation) Let ϕ be a representation of G on V. For V, we can define its dual space, $V^* = \mathrm{Hom}(V, \mathbb{C})$. We define $\phi^* : G \to GL(V^*)$ by $\phi^*(g) = \phi^t(g^{-1})$ for all $g \in G$. ϕ^* is the **dual representation**. It is easy to check that this is actually a representation: let $g, h \in G$ and ρ_g, ρ_h be their associated mappings in $GL(V)$. Then,

$$\begin{aligned} \rho^*_{gh} &= \rho^t_{(gh)^{-1}} \\ &= (\rho_{h^{-1}} \circ \rho_{g^{-1}})^t \\ &= \rho^t_{g^{-1}} \circ \rho^t_{h^{-1}} \\ &= \rho^*_g \circ \rho^*_h \end{aligned}$$

as we wanted.

Examples 2.7. If V, W are representations of G, the direct sum $V \oplus W$ and the tensor product $V \otimes W$ are also representations, the latter via

$$g(v \otimes w) = gv \otimes gw.$$

For a representation V, the n^{th} tensor power $V^{\otimes n}$ is again a representation of G by this rule, and exterior powers $\Lambda^n(V)$ and symmetric powers $\mathrm{Sym}^n(V)$ are subrepresentations of it.

3. Lie Group

Here, we introduce concept of Lie group, which plays crucial role in physics, particularly in studies of particle physics. We make a slight detour to introduce an application in physics and the necessary concepts in topology.

Definition 3.1. A **real Lie group** is a group that is also a finite-dimensional real smooth manifold, in which the group operations of multiplication and inversion are smooth maps. Smoothness of the group multiplication

$$\mu : \; G \times G \to G, \; \mu(x, y) = xy$$

means that μ is a smooth mapping of the product $G \times G$ into G. These two requirements can be combined to the single requirement that the mapping

$$(x, y) \mapsto x^{-1}y$$

be a smooth mapping of $G \times G$ into G.

As defined above, Lie group embodies three different forms of mathematical structure. Firstly, it has the group structure. Secondly, the elements of this group also form a "topological space" so that it may be described as being a special case of a "topological group". Finally, the elements also constitute an "analytic manifold". Consequently, a Lie group may be defined in several different (but equivalent) ways, depending on degree of emphasis on its various aspects. In particular, it can be defined as a topological group with certain analytic properties or, alternatively, as an analytic manifold with group properties. But, formulating in these ways would require many set of other definitions (such as manifolds, smooth mapping, and etc.), which may not be very important in understanding the applications of Lie groups in physics. In fact, we are mainly interested in a particular type of Lie group for problems in physics, the matrix Lie group, for which a straightforward definition can be given using the general linear group, $GL(n;\mathbb{C})$.

Definition 3.2. The **general linear group** over the real numbers, denoted by $GL(n;\mathbb{R})$ is the group of all $n \times n$ invertible matrices with real number entries. We can similarly define it over the complex numbers, $\mathbb{C}$ denoted by $GL(n;\mathbb{C})$.

Definition 3.3. A **matrix Lie group** is any subgroup H of $GL(n;\mathbb{C})$ with the following property: if A_n is any sequence of matrices in H, and A_n converges to some matrix A, then either $A \in H$, or A is not invertible. The condition on H amounts to saying that H is a closed subset of $GL(n;\mathbb{C})$. Thus, one can think matrix Lie group as simply a **closed** subgroup of $GL(n;\mathbb{C})$.

Example 3.4. (Counterexamples) An example of a subgroup of $GL(n;\mathbb{C})$ which is not closed is the set of all $n \times n$ invertible matrices whose entries are real and rational numbers. One can easily have a sequence of invertible matrices with rational number entries converging to an invertible matrix with some irrational number entries.

Example 3.5. (The general linear groups, $GL(n;\mathbb{C})$ and $GL(n;\mathbb{R})$) The **general linear groups** are themselves matrix Lie groups. Of course, $GL(n;\mathbb{C})$ is a subgroup of itself. Also, if A_n is a sequence of matrices in $GL(n;\mathbb{C})$ and A_n converges to A, then by definition of $GL(n;\mathbb{C})$, either A is in $GL(n;\mathbb{C})$ or A is not invertible.
Moreover, $GL(n;\mathbb{R})$ is a subgroup of $GL(n;\mathbb{C})$ and if $A_n \in GL(n;\mathbb{R})$ and A_n converges to A, then the entries of A are, of course, real. Thus, either A is not invertible, or $A \in GL(n;\mathbb{R})$.

Example 3.6. (The special linear groups $SL(n;\mathbb{C})$ and $SL(n;\mathbb{R})$) The **special linear groups** is the group of $n \times n$ invertible matrices having determinant 1. Since determinant is a continuous function, if a sequence A_n in $SL(n;\mathbb{C})$ converges to A, then A also has a determinant 1 and $A \in SL(n;\mathbb{C})$

Example 3.7. (The orthogonal and special orthogonal groups, $O(n)$ and $SO(n)$) An $n \times n$ matrix A is **orthogonal** if the column vectors that make up A are orthonormal, that is, if

$$\sum_{i=1}^{n} A_{ij} A_{ik} = \delta_{jk}$$

Equivalently, A is orthogonal if it preserves inner product, namely if

$$\langle x, y \rangle = \langle Ax, Ay \rangle$$

for all $x, y \in \mathbb{R}^n$. Another equivalent definition is that A is orthogonal if $A^t A = \mathbb{I}$, i.e. if $A^t = A^{-1}$. Since $\det A^t = \det A$, if A is orthogonal, then $\det A = \pm 1$. Hence, orthogonal matrix must be invertible. Furthermore, if A is an orthogonal matrix, then

$$\langle A^{-1}x, A^{-1}y\rangle = \langle A(A^{-1}x), A(A^{-1}y)\rangle = \langle x, y\rangle.$$

Thus, the inverse is also orthogonal. Also, the product of two orthogonal matrices is orthogonal. Therefore, the set of $n \times n$ orthogonal matrices forms a group, called **orthogonal group** and it is a subgroup of $GL(n;\mathbb{C})$. The limit of a sequence of orthogonal matrices is orthogonal since the relation $A^t A = \mathbb{I}$ is preserved under limits. Thus, $O(n)$ is a matrix Lie group.
Similar to $SL(n;\mathbb{C})$, the **special orthogonal group**, denoted by $SO(n)$, is defined as subgroup of $O(n)$ whose matrices have determinant 1. Again, this is a matrix Lie group.

Remark 3.8. Geometrically speaking, the elements of $O(n)$ are either rotations, or combinations of rotations and reflections, while the elements of $SO(n)$ are just the rotations. Due to this geometric nature, the special orthogonal group appears frequently in physics problem dealing with a rotational symmetry. An example of this would be a problem dealing with hydrogen atom potentials, which has a spherical symmetry.

Example 3.9. (The unitary and special unitary groups, $U(n)$ and $SU(n)$) An $n \times n$ complex matrix A is **unitary** if the column vectors of A are orthonormal that is, if

$$\sum_{i=1}^{n} \overline{A_{ij}} A_{ik} = \delta_{jk}$$

Similar to an orthogonal matrix, a unitary matrix has two another equivalent definitions. A matrix A is unitary,
(1) If it preserves an inner product.
(2) If $A^*A = 1$, i.e. if $A^* = A^{-1}$ (where A^* is adjoint of A)
Since $\det A^* = \overline{\det A}$, $|\det A| = 1$ for all unitary matrices A. This shows unitary matrices are invertible. The same argument as for the orthogonal group can used to show that the set of unitary matrices form a group, called **unitary group** $U(n)$. This is clearly a subgroup of $GL(n;\mathbb{C})$ and since limit of unitary matrices is unitary, $U(n)$ is a matrix Lie group. The subgroup of unitary group whose matrices have determinant 1 is the **special unitary group** $SU(n)$. It is easy to see that this is also a matrix Lie group.

Here, we take a short detour to cover some needed topics in topology to help understand an application of special unitary group, $SU(2)$ and special orthogonal group, $SO(3)$, in physics.

Definition 3.10. A **covering space** of a space X is space $\tilde{X}$ together with a map $\rho: \tilde{X} \to X$, satisfying the following condition: There exists an open cover $\{U_\alpha\}$ of X such that, for each α, $\rho^{-1}(U_\alpha)$ is a disjoint union of open sets in $\tilde{X}$. Likewise, **covering group** can be defined similarly on topological groups, in particular matrix Lie groups.

Definitions 3.11. A **path** in topological space X is a continuous function f from unit interval $I = [0, 1]$ to X

$$f : I \to X$$

A space X is said to be **path-connected** if any two points on X can be joined by a path. The stronger notion, the **simply-connected** space X, is if:
(1) X is path-connected,
(2) and every path between two points can be continuously transformed, staying within space, into any other such path while preserving two endpoints.

Definition 3.12. A covering space is a **universal covering space** if it is simply-connected. The name *universal cover* comes from the property that the universal cover (of the space X) covers any connected cover (of the space X), i.e. if the mapping $\rho:\ \overline{X} \to X$ is the universal cover of the space X and the mapping $\phi:\ \tilde{X}$ is any cover of the space X where $\tilde{X}$ is connected, then there exists a covering map $\psi:\ \overline{X} \to \tilde{X}$ such that $\phi \circ \psi = \rho$.

Proposition 3.13. *The matrix Lie group $SU(2)$ can be identified with the manifold S^3.*

Proof. Consider
$$U = \begin{pmatrix} \alpha & \mu \\ \beta & \nu \end{pmatrix} \in SU(2)$$
for $\alpha, \beta, \mu, \nu \in \mathbb{C}$. This is in $SU(2)$ if
$$U^*U = \mathbb{I}$$
and
$$\det U = 1.$$
Using the inverse matrix formula with $\det U = 1$,
$$U^{-1} = \begin{pmatrix} \nu & -\mu \\ -\beta & \alpha \end{pmatrix}$$
Since $U^{-1} = U^*$, we have
$$\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}$$
where
$$U = \alpha\overline{\alpha} + \overline{\beta}\beta = 1$$
This is a generic form of elements of $SU(2)$. Now, set $\alpha = y_0 + iy_3, \beta = -y_2 + iy_1$ for $y_0, y_1, y_2, y_3 \in \mathbb{R}$. Then, it is straightforward to see that
$$U = y_0\mathbb{I} + \sum_{n=1}^{3} y_n\sigma_n$$
where σ_n are Pauli matrices defined as following:
$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
Now, the previous condition $\alpha\overline{\alpha} + \overline{\beta}\beta = 1$ is equivalent to $y_0^2 + y_1^2 + y_2^2 + y_3^2 = 1$ i.e. $y = (y_0, y_1, y_2, y_3) \in S^3$. This establishes a smooth, invertible map between $SU(2)$ and S^3. □

Remark 3.14. The Pauli matrices,$\{\sigma_n\}$, which are familiar from Quantum mechanics, can be taken as generators for Lie algebra of $SU(2)$, often with an extra factor of i, in physics: $\{i\sigma_n\}$.

Remark 3.15. The point of establishing the correspondence between $SU(2)$ and S^3 is to provide an explicit way of seeing that $SU(2)$ is simply-connected, since unit sphere S^n (with $n \geq 2$) is simply-connected, in particular, S^3.

Proposition 3.16. *The special unitary group, $SU(2)$ is a double-cover of the special orthogonal group $SO(3)$. There is 2-1 correspondence $\phi : SU(2) \to SO(3)$ and ϕ is a group homomorphism. Additionally, it is a universal cover.*

Proof. Suppose $U \in SU(2)$. Define a 3×3 matrix, $\phi(U)$ via

$$\phi(U)_{mn} = \frac{1}{3}\text{trace}(\sigma_m U \sigma_n U^*),$$

where σ_n are Pauli matrices defined earlier. By writing $U = y_0\mathbb{I} + \sum_{n=1}^{3} y_n\sigma_n$ for $y_0, y_n \in \mathbb{R}$, satisfying ${y_0}^2 + \sum_{p=1}^{3} y_p y_p = 1$. It is straightforward from here to show that

$$\phi(U)_{mn} = ({y_0}^2 - \sum_{p=1}^{3} y_p y_p)\ \delta_{mn} + 2\ \epsilon_{mnq} \sum_{q=1}^{3} y_0 y_q + 2\ y_m y_n$$

It is clear that if $y_p = 0$ for $p = 1, 2, 3$, so that $U = \pm\mathbb{I}$, then $\phi(U) = \mathbb{I}$, so $\phi(U) \in SO(3)$. More generally, suppose that $y_p \neq 0$, then we can set $y_0 = \sin(\alpha), y_p = \cos(\alpha)z_p$ for $0 < \alpha < 2\pi$, $\alpha \neq \pi$. Then the constraint ${y_0}^2 + \sum_{p=1}^{3} y_p y_p = 1$ implies that $\sum_{p=1}^{3} z_p z_p = 1$, i.e. $\vec{z} := (z_1, z_2, z_3)$ is a unit vector in $\mathbb{R}^3$. Then $\phi(U)_{mn}$ can be rewritten as following:

$$\phi(U)_{mn} = \cos(2\alpha)\delta_{mn} + \sum_{q=1}^{3} \sin(2\alpha)\epsilon_{mnq} z_q + (1 - \cos(2\alpha)) z_m z_n.$$

It is then apparent that

$$\phi(U)_{mn} z_n = z_m,$$

and if $\vec{x}$ is orthogonal to $\vec{z}$ then

$$\phi(U)_{mn} x_n = \cos(2\alpha) x_m + \sum_{q=1}^{3} \sin(2\alpha)\epsilon_{mnq} x_n z_q$$

The transformation ϕ is therefore corresponds to a rotation by 2α in the plane with unit normal vector $\vec{z}$.
It is clear that any non-trivial rotation in $SO(3)$ can be written in this way. However, the correspondence is not 1 - 1, but 2-1. To see this explicitly, it requires bit of tedious algebras; so we will skip this part of the proof. The end picture will be that $\phi(U)$ from $SO(3)$ will correspond to both U and $-U$ from $SU(2)$.
Now, check the group homomorphism, i.e., $\phi(U_1 U_2) = \phi(U_1)\phi(U_2)$ for $U_1, U_2 \in SU(2)$. Let us write U_1, U_2 by following: (Note that we are now using the Einstein notation, where there is implicitly a summation over repeated indices.)

$$U_1 = y_0\mathbb{I} + iy_n\sigma_n,\ U_2 = w_0\mathbb{I} + iw_n\sigma_n$$

for $y_0, y_p, w_0, w_p \in \mathbb{R}$ satisfying $y_0^2 + y_p y_p = w_0^2 + w_p w_p = 1$ then

$$U_1 U_2 = u_0\mathbb{I} + iu_n\sigma_n$$

where $u_0 = y_0 w_0 - y_p w_p$ and $u_m = y_0 w_m + y_m w_0 - \epsilon_{mpq} y_p w_q$ satisfy $u_0^2 + u_p u_p = 1$. It then suffices to evaluate directly

$$\phi(U_1 U_2)_{mn} = (u_0^2 - u_p u_p)\delta_{mn} + 2\epsilon_{mnq} u_0 u_q + 2u_m u_n$$

and compare this with

$\phi(U_1)_{mp}\phi(U_2)_{pn} = [(y_0^2 - y_l y_l)\delta_{mp} + 2\epsilon_{mpq} y_0 y_q + 2y_m y_p][(w_0^2 - w_r w_r)\delta_{pn} + 2\epsilon_{pnr} w_0 w_k + 2w_p w_n]$

From here, it is, again, a simple, but tedious matter of expanding out two expressions in terms of y and w and checking $\phi(U_1U_2)_{mn} = \phi(U_1)_{mp}\phi(U_2)_{pn}$ as required.

The final statement about the universal cover follows straightforward from simple-connectedness of $SU(2)$ and the definition of universal cover. □

Remark 3.17. To quote from W. S. Massey, *Algebraic Topology: An Introduction*, "A simply-connected space admits no non-trivial coverings." Equivalently speaking, a universal cover is unique up to a homeomorphism. In linear group (or matrix Lie group), in particular, the only cover of simply-connected group is an isomorphism; thus, the universal covers are isomorphic. This fact, in addition, to the relationship between $SU(2)$ and $SO(3)$ has a well-known, physical outcome: There are particles with only integer or half-integer spins (bosons or fermions), i.e. there is no 1/3 or $1/N$ ($N \neq 2$) spin particles.

4. Lie Algebra

Now, we move on to Lie algebras. The concept of Lie algebra can be motivated as the tangent space of the associated Lie group at the identity, using Lie group's smooth manifold structure. Instead, we can also define Lie algebras, using the matrix exponential, which is much more straightforward. First, we define the matrix exponential.

Definition 4.1. Let X be any $n \times n$ complex matrix. We define the exponential of X, e^X, by the usual power series:

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}.$$

Proposition 4.2. *For any $n \times n$ complex matrix X, the series above converges. The matrix exponential, e^X is a continuous function of X.*

Proof. Recall that the norm of matrix A is defined as

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Now, we can think of an $n \times n$ complex matrix X as residing in $\mathbb{C}^{n^2}$, which is a complete space. Using the norm property from earlier, it is easy to note that

$$\|X^m\| \leq \|X\|^m,$$

and hence

$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \leq \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} < \infty.$$

Thus, the exponential series converges absolutely. But, given absolute convergence, we can take partial sums of the series to form a Cauchy sequence. Now, using the fact that every Cauchy sequence converges in a complete space, we now have that the exponential series converges.

As for continuity, we just have to note that X^m is a continuous function of X. □

Proposition 4.3. *Let X, Y be arbitrary $n \times n$ complex matrices. Then following are true:*
1) $e^0 = \mathbb{I}$.
2) e^X *is invertible and* $(e^X)^{-1} = e^{-X}$.
3) $e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ *for all complex numbers* α, β.
4) If $XY = YX$, *then* $e^{X+Y} = e^X e^Y = e^Y e^X$.
5) If C *is an* $n \times n$ *invertible complex matrix, then* $e^{CXC^{-1}} = Ce^X C^{-1}$.
6) $\|e^X\| \leq e^{\|X\|}$.

Proof. Point 1) is obvious from the fact that $X^0 = I$ for all matrices X and $0^m = 0$ for all $m \geq 1$. Point 2) and 3) are special cases of 4). To verify 4), we simply multiply power series out term by term.

$$e^X e^Y = \left(I + X + \frac{X^2}{2!} + \cdots \right) \left(I + Y + \frac{Y^2}{2!} + \cdots \right).$$

By collecting the terms where the power of X and the power of Y add up to m, we get

$$e^X e^Y = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k}.$$

Since we are given $XY = YX$, we have

$$(X+Y)^m = \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k}.$$

Substituting this back in to what we had earlier, we get

$$e^X e^Y = \sum_{m=0}^{\infty} \frac{1}{m!} (X+Y)^m = e^{X+Y}.$$

For 5) we simply note that, $(CXC^{-1})^m = CX^m C^{-1}$. The proof of 6) was already made from the proof of proposition 3.19. □

Definition 4.4. Let G be a matrix Lie group. Then the **Lie algebra** of G, denoted $\mathfrak{g}$, is the set of all matrices X such that e^{tX} is in G for all real number t, together with a bracket operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the **Lie bracket**, with the following properties:
1) $[\cdot, \cdot]$ is anti-symmetric, i.e. for all $X, Y \in \mathfrak{g}$,

$$[X, Y] = -[Y, X]$$

2) $[\cdot, \cdot]$ is bilinear, i.e. for all $a, b \in \mathbb{C}$ and $X, Y, Z \in \mathfrak{g}$,

$$[aX + bY, Z] = a[X, Z] + b[Y, Z]$$

and

$$[X, aY + bZ] = a[X, Y] + b[X, Z]$$

3) $[\cdot, \cdot]$ satisfies the Jacobi identity, i.e. for all $X, Y, Z \in \mathfrak{g}$,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

In fact, for a matrix Lie group G, the Lie bracket associated with its Lie algebra $\mathfrak{g}$

is simply given by commutator of matrices, i.e. for all $X, Y \in \mathfrak{g}$, $[X, Y]$ is defined by:

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

and it is easy to check that three properties metioned above are satisfied.

Remark 4.5. Physicists are accustomed to considering the map $X \to e^{iX}$ instead of usual $X \to e^X$. Thus, in physics, Lie algebra of G is often defined as set of matrices X such that $e^{itX} \in G$ for all real t. In physics, Lie algebra is frequently referred to as the space of "infinitesimal group elements", which actually connects the concept of Lie algebra back to its original definition as the tangent space.

Example 4.6. (The general linear groups) Let X be any $n \times n$ complex matrix, then by proposition 3.20, e^{tX} is invertible. Thus, the Lie algebra of $GL(n; \mathbb{C})$ is the space of all $n \times n$ complex matrices. This Lie algebra is denoted by $\mathfrak{gl}(n; \mathbb{C})$.
If X is any $n \times n$ real matrix, then e^{tX} will be invertible and real. On the other hand, if e^{tX} is real for all t, then $X = \frac{d}{dt} |_{t=0}\, e^{tX}$ will also be real. Thus the Lie algebra of $GL(n; \mathbb{R})$ is the space of all $n \times n$ real matrices, denoted $\mathfrak{gl}(n; \mathbb{R})$.

Proposition 4.7. *Let X be $n \times n$ complex matrix. Then,*

$$\det(e^X) = e^{\text{trace}(x)}$$

Proof. We can divide up the proof into three cases: X is diagonalizable, nilpotent, or arbitrary. The reason we can do this is because any matrix X can be written in the form: $X = S + N$ with S diagonalizable, N nilpotent, and $SN = NS$. This follows from the Jordan canonical form. Since S and N commute,

$$e^{S+N} = e^S e^N$$

and we can, then, use the results for diagonalizable and nilpotent matrices to compute for the arbitrary matrices.
Case 1: Suppose X is diagonalizable. Then, there exists an inveritble matrix C such that

$$X = C \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} C^{-1}.$$

Then

$$e^X = C \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} C^{-1},$$

using the proposition 3.20. Thus, $\text{trace}(X) = \sum \lambda_i$, and $\det(e^X) = \prod e^{\lambda_i} = e^{\sum \lambda_i}$.
Case 2: Suppose X is nilpotent. If X is nilpotent, then it cannot have any non-zero eigenvalues. Thus, all the roots of characteristic polynomial must be zero. Hence, the Jordan canonical form of X will be strictly upper-triangular. X can be written as:

$$X = C \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} C^{-1}.$$

Hence, e^X will be upper-triangular with 1's on the diagonals:

$$e^X = C \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} C^{-1}.$$

Thus, if X is nilpotent, $\text{trace}(X) = 0$ and $\det e^X = 1$.
Case 3: let X be arbitrary. Then, $e^X = e^S e^N$ as described above.

$$\det e^X = \det e^S \det e^N = e^{\text{trace}(S)} e^{\text{trace(N)}} = e^{\text{trace}(X)},$$

as we wanted to show. □

Example 4.8. (The special linear groups) We have $\det(e^X) = e^{\text{trace}X}$. Thus, if $\text{trace}X = 0$, then $\det(e^{tX}) = 1$ for all real number t. On the other hand, if X is any $n \times n$ matrix X such that $\det(e^{tX}) = 1$ for all real number t, then $e^{(t)(\text{trace}X)} = 1$ for all t. This means that $(t)(\text{trace}X)$ is an integer multiple of $2\pi i$ for all t, which is only possible if $\text{trace}X = 0$. Thus the Lie algebra of $SL(n; \mathbb{C})$ is the space of all $n \times n$ complex matrix with trace zero, denoted $\mathfrak{sl}(n; \mathbb{C})$.
Similarly, the Lie algebra of $SL(n; \mathbb{R})$ is the space of all $n \times n$ real matrices with trace zero, denoted $\mathfrak{sl}(n; \mathbb{R})$.

Example 4.9. (The unitary groups) Recall that matrix U is unitary if and only if $U^* = U^{-1}$. Thus, e^{tX} is unitary if and only if

$$(e^{tX})^* = (e^{tX})^{-1} = e^{-tX}.$$

But by taking adjoints term-by-term, we see that $(e^{tX})^* = e^{tX^*}$, and so the above becomes

$$e^{tX^*} = e^{-tX}.$$

Clearly, a sufficient condition is $X^* = -X$. On the other hand, if above hold for all real t, then by differentiating at $t = 0$, we see that $X^* = -X$ is necessary condition. Thus the Lie algebra of $U(n)$ is the space of all $n \times n$ complex matrices X such that $X^* = -X$, denoted $\mathfrak{u}(n)$. By combining the result for the special linear groups, we see that Lie algebra of $SU(n)$ is the space of all $n \times n$ complex matrices X such that $X^* = -X$ and $\text{trace}X = 0$, denoted $\mathfrak{su}(n)$.

Proposition 4.10. *Let G be a matrix Lie group, and X and element of its Lie algebra. Then e^X is an element of identity component of G. The **identity component** of a topological group G is the connected component G_0 of G that contains the identity element of the group.*

Proof. By defintion of Lie algebra, e^{tX} lies in G for all real t. But as t varies from 0 to 1, e^{tX} is a continuous path connecting identity to e^X. □

Example 4.11. (The orthogonal groups) The identity component of $O(n)$ is just $SO(n)$. By the proposition 3.27, the exponential of a matrix in Lie algebra is automatically in the identity component. So, the Lie algebra of $O(n)$ is the same as the Lie algebra of $SO(n)$.
Now, an $n \times n$ real matrix X is orthogonal if and only if $X^{tr} = X^{-1}$. (Note, here, we used X^{tr} instead of X^t to not cause confusion with the t for the exponential.) So, given an $n \times n$ real matrix X, e^{tX} is orthogonal if and only if $(e^{tX})^{tr} = (e^{tX})^{-1}$, or

$$e^{tX^{tr}} = e^{-tX}$$

Clearly, a sufficient condition for above to hold is that $X^{tr} = -X$. Meanwhile, if above equality holds for all t, then by differentiating at $t = 0$, we must have $X^{tr} = -X$. Thus the Lie algebra of $O(n)$, as well as $SO(n)$, is the space of all $n \times n$ real matrices X with $X^{tr} = -X$, denoted $\mathfrak{so}(n)$. Note that the condition $X^{tr} = -X$ forces the diagonal entries of X to be zero, and so explicitly the trace of X is zero.

5. Physical Applications

Here, we look at how the concepts introduced earlier have been applied in the field of physics, particularly, particle physics.

5.1. **Isospin and $SU(2)$.**

The simplest case of an application in physics can be found in a Lie algebra generated from the bilinear products of creation and annihilation operators where there are only two quantum states. This is often referred to as the "old fashioned" isospin as it was originally conceived for systems of neutrons and protons before the discovery of mesons and strange particles. The concept of isospin was first introduced by Heisenberg in 1932 to explain the symmetries of newly discovered neutrons. Although the proton has a positive charge, and neutron is neutral, they are almost identical in other respects such as their masses. Hence, the term 'nucleon' was coined: treating two particles as two different states of the same particle, the nucleon. In fact, the strength of strong interaction - the force which is responsible forming the nucleus of an atom - between any pair of nucleons is independent of whether they are interacting as protons and nucleons. More precisely, the isospin symmetry is given by the invariance of Hamiltonian of the strong interactions under the action of Lie group, $SU(2)$. The neutrons and protons are assigned to the doublets with spin 1/2-representation of $SU(2)$. Let us take a more detailed look in the mathematical formulation.

Let $a_p^\dagger$ and $a_n^\dagger$ be operators for the creation of a proton and neutron, respectively, and let a_p and a_n be the corresponding annihilation operators. Now, construct the Lie algebra of all possible bilinear products of these operators which do not change the number of particles (strong interaction invariance). There are four possible bilinear products:

$$a_p^\dagger a_n, \quad a_n^\dagger a_p, \quad a_p^\dagger a_p, \quad a_p^\dagger a_p$$

The first operator turns a neutron into a proton, while the second operator turns a proton into a neutron. Let us denote the first two operators by τ_+ and τ_-. Recall, this whole symmetry is based on the idea that the proton and the neutron are simply two different states of the same particle: we can treat the proton as having spin-up and the neutron as having spin-down, i.e., associating them with doublets $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. With that in mind, the notation would seem more natural: τ_+ being the raising operator, while τ_- being the lowering operator. Now, the last two operators simply annihilates a proton or neutron and then create them back. These are just the number operators which count the number of protons and neutrons. Together, they are total number operators, which commute with the all the other operators, as they do not change the total number. It is therefore convenient to divide the set of four operators into a set of three operators plus the total number, or Baryon number, operator, which commutes with others:

$$
\begin{aligned}
B &= a_p^\dagger a_p + a_n^\dagger a_n, \\
\tau_+ &= a_p^\dagger a_n \\
\tau_- &= a_n^\dagger a_p \\
\tau_0 &= \frac{1}{2}(a_p^\dagger a_p - a_n^\dagger a_n) = Q - \frac{1}{2}B
\end{aligned}
$$

where Q is just the total charge (Since proton has a positive charge whereas neutron has no charge.). Now, the set of three operators, τ_+, τ_-, τ_0 satisfy following commutation relation, which is exactly like that of angular momenta:

$$
\begin{aligned}
[\tau_0, \tau_+] &= \tau_+ \\
[\tau_0, \tau_-] &= -\tau_- \\
[\tau_+, \tau_-] &= 2\tau_0
\end{aligned}
$$

This has led to the designation isospin for these operators and to the description of rotations in the fictitious isospin space.

Let us now consider which Lie group is associated with these isospin operators. By analogy with angular momentum operators, we allow these operators to generate infinitesimal transformations such as

$$\psi' = \{1 + i\epsilon(\tau_+ + \tau_-)\}\psi.$$

We use the linear combination $\tau_+ + \tau_-$ because these are not individually Hermitian (or self-adjoint). Note that such a transformation changes a proton or neutron into something which is a linear combination of the proton and the neutron state. These transformations are thus transformations in a two-dimensional proton-neutron Hilbert space. The transformations are unitary thus the Lie group associated with isospin is some group of unitary transformation in a two-dimensional space. The whole group of unitary transformations in a two-dimensional space is generated by the set of four operators; however, the unitary transformations generated by the operator B are of a trivial nature: they are multiplication of any state by a phase factor. Since the three isospin operators form a Lie group by themselves, the associated group is the subgroup of full unitary group in two-dimensions, the special unitary group, $SU(2)$.

5.2. **Eightfold way and $SU(3)$.** As noted before, the above isospin $SU(2)$ symmetry is old fashioned in that it does not consider mesons and the "strangeness" - a property in particles, expressed as quantum numbers, for describing a decay of particles in strong and electromagnetic interactions, which occur in a short amount of time. This was first introduced by Murray Gell-Mann and Kazuhiko Nijishima to explain that certain particles such as the kaons or certain hyperons were created easily in particle collisions, yet decayed much more slowly than expected for their large masses. To account for the newly added property (quantum number), the Lie group $SU(3)$ was chosen over $SU(2)$ to construct a theory which organizes baryons and mesons into octets (thus, the term Eightfold Way), where the octets are the representations of the Lie group, $SU(3)$.

Acknowledgments. I would like to thank the University of Chicago Math Department for funding this program, and Professor May for running it. Special thanks are also due to Henry Chan and Jonathan Rubin for their guidance and help during the program and the writing of this paper.

References

[1] Woit, Peter. *Quantum Theory, Groups and Representations: An Introduction.* http://www.math.columbia.edu/ woit/QM/qmbook.pdf
[2] Fulton, William and Joe Harris *Representation Theory.* Springer, 2004
[3] Cornwell, John F. *Group Theory in Physics.* Academic Press. 1997
[4] Sattinger, D.H. and O.L. Weaver. *Lie Groups and Algeabras with Application to Physics, Geometry, and Mechanics.* Springer-Verlag, 1986.
[5] Lipkin, Harry J. *Lie Groups for Pedestrians.* Dover Publications, INC. 1966