A Note on Group Theory

© *Chengchao Yuan*[*], AST@NJU(12/2015)

§. Preface

This note provides a comprehensive approach to the general ideas of group theory and the knowledge of differential geometry such as the calculus on manifolds is not required. Section 1 and section 2 are primarily taken from the textbook *An introduction to tensors and group theory* while the most representative examples such as $SU(2), GL(n,\mathbb{R}), U(n)$ are preserved. The Lie groups and Lie algebra are discussed in a way that is easy to understand without abstract mathematic definitions. Further applications to theoretical physics especially in Quantum Field Theory are introduced in section 4 where we use Lie algebra to deduce the single-particle relativistic wave equation (also known as Dirac equation). As I understand, if you can differentiate the physical meanings (rotation or boost) of $SU(2), SO(3)$ and Lorentz group as well as their corresponding Lie algebras, the goal of this proof-not-so-much note is almost achieved.

§. Definition and Examples

A group is a set G together with a "multiplication" operation, denoted $\cdot$, that satisfies the following four axioms:

1. (Closure) $g, h \in G$ implies $h \cdot g \in G$.

2. (Associativity) For $h, k, g \in G$, $g \cdot (h \cdot k) = (g \cdot h) \cdot k$.

3. (Existence of identity). there exists an element $e \in G$ such that $g \cdot e = e \cdot g = g$, $\forall g \in G$.

4. (Existence of reverse) $\forall g \in G$, there exist an element $h \in G$ such that $g \cdot h = h \cdot g = e$ and h is denoted $h = g^{-1}$.

Some properties:

$$\begin{aligned} (g^{-1})^{-1} &= g \\ (g \cdot h)^{-1} &= h^{-1} \cdot g^{-1} \\ e^{-1} &= e \end{aligned} \tag{1}$$

We may think the group as a set of transformations. Consider two arbitrary elements of a group $g, h \in G$, if $g \cdot h = h \cdot g$, G is called to be commutative(or abelian). If there exist two elements, $g, h \in G$, satisfy $g \cdot h \neq h \cdot g$, G is called non-commutative(or non-abelian).

Example 1 $\mathbb{R}$: The real numbers as an addictive group.

The group "multiplication" operation is given by regular addition, i.e.,

$$x \cdot y = x + y \tag{2}$$

Example 2 $\mathbb{R}/\{0\}$ is an abelian group under regular multiplication.

Example 3 $GL(V), GL(n,\mathbb{R})$ and $GL(n,\mathbb{C})$: The general linear groups.

[*]Email: yuancch@outlook.com

The general linear group of a vector space V, denoted $GL(V)$, is defined to be the subset of $\mathcal{L}(V)$ consisting all *invertible* linear operators on V while $\mathcal{L}(V)$ is a vector space formed by *all* linear operators on V. If V have a scalar field $\mathbb{C}$ and dimension n. For each $T \in GL(V)$, we get an invertible matrix $[T] \in M_n(\mathbb{C})$. Just as the invertible $T \in \mathcal{L}(V)$ form a group, so do the corresponding invertible matrices in $M_n(\mathbb{C})$; this group is denoted $GL(n, \mathbb{C})$, named the complex general linear group in n dimensions. Similarly we can defind $GL(n, \mathbb{R})$.

Example 4 Isometries

Suppose we have a vector space equipped with a non-degenerate Hermitian form $\langle \cdot | \cdot \rangle$. In this case we can consider the set of **isometrics** $\mathrm{Isom}(V)$ consisting of those operators T with "preserve" $\langle \cdot | \cdot \rangle$ in the sense that

$$\boxed{\langle Tv | Tw \rangle = \langle v | w \rangle \,, \ \forall v, w \in V.} \tag{3}$$

An isometry T can be thought of as an operator that preserve the "lengths" of vectors since "iso-metry" equals to "same length". It is easy to prove that $\mathrm{Isom}(V)$ is a subgroup of $GL(V)$.

Example 5 The orthogonal group $O(n)$

Let V be the n-dimensional real inner product space. The isometries of V can be thought of an operators that preserve the length and angle, since the formula

$$\cos\theta = \frac{\langle v | w \rangle}{\|v\| \|w\|} \tag{4}$$

Now if T is an isometry, we find that

$$\begin{aligned} [v]^T [w] &= \langle v | w \rangle \\ &= \langle Tv | Tw \rangle \\ &= \delta_{ij} (Tv)^i (Tw)^j \\ &= \sum_i v^k T^i_k T^i_l w^l \\ &= [v]^T [T]^T [T] [w], \ \forall v, w \in V \end{aligned} \tag{5}$$

This is true only if

$$[T]^T = [T]^{-1} \tag{6}$$

The set of all orthogonal matrices is called orthogonal group $O(n)$.

$SO(3)$ group is strongly related to the rotation in 3-dimension space. An arbitrary rotation can be described by three Euler angles: ψ, θ, ϕ and together give a general form for $R \in SO(3)$:

$$\begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

.

Example 6 The unity group $U(n)$

Now let V be a complex inner product space. First we define the adjoint* $T^\dagger$ of a linear operator T by the equation

$$\boxed{\langle T^\dagger v|w\rangle = \langle v|Tw\rangle} \tag{8}$$

If T is an isometry, then we can characterize it in terms of its joint, as follow: first we have

$$\langle v|w\rangle = \langle Tv|Tw\rangle = \langle T^\dagger Tv|w\rangle\ ,\ \forall v, w \in V. \tag{9}$$

This is equivalent to $T^\dagger = T^{-1}$. Such an operator is called *unitary*. Every isometry of a complex inner product space is unitary and vice versa. In an orthonormal basis, a unitary operator is represented by a unitary matrix. Thus, $[T^\dagger] = [T]^\dagger = [T]^{-1}$. A set of all unitary matrices form a group denoted $U(2)$.

Example 7 The Lorentz group $O(n-1,1)$

Now let V be a real vector space with $Minkowskimetric \eta$, which is defined as a symmetric, non-degenerate $(2,0)$ tensor whose matrix in an orthonormal basis has form

$$[\eta] = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \dots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix}. \tag{10}$$

Since η is non-degenerate Hermitian form $\eta(v,w) = [v]^T[\eta][w]$, we can consider its group of isometries. If $T \in$Isom(V), we have

$$[v]^T[\eta][w] = \eta(v,w) = \eta(Tv,Tw) = [v]^T[T]^T[\eta][T][w], \tag{11}$$

which leads to $[T]^T[\eta][T] = [\eta]$ or in the components reads $T^\rho_\mu T^\sigma_\nu \eta_{\rho\sigma} = \eta_{\mu\nu}$. In fact this is the definition of Lorentz transformation. The set of all Lorentz transformations forms a group, known as Lorentz group.

$SO(3,1)$, the restricted Lorentz group, is defined to be the set of all $A \in O(3,1)$ which satisfy $|A| = 1$ as well as $A_{44} > 1$. Such transformations are known as restricted Lorentz transformation. The famous transformation is

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -\beta\gamma \\ 0 & 0 & -\beta\gamma & \gamma \end{pmatrix}, -1 < \beta < 1, \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{12}$$

. Such a transformation is also known as a boost along the z axis.

Example 8 $SU(n)$ **and** $SO(n)$

Special unitary group $SU(n)$; Special orthogonal group $SO(n)$.

---

*Transpose: Consider a linear operator A on a vector space V. We can define a linear operator on V^* called the transpose of A and denoted by A^T as follows $(A^T(f))(v) = f(Av)$ where $v \in V, f \in V^*$. L is the map from V to V^*. Now we can construct $A^\dagger$; this is known as Hermitian adjoint of A, and is defined as $A^\dagger = L^{-1} \circ A^T \circ L : V \to V$.

$SU(2)$ is the group of special unity group in two complex dimensional matrices A which satisfy $|A| = 1$ and $A^\dagger = A^{-1}$. A general element of $SU(2)$ looks like

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1. \tag{13}$$

Example 9 S_n: The symmetric group in n letters

Any permutation σ is specified by n numbers $\sigma(i), i = 1, \ldots, n$ and can be conveniently be be notated as

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} \tag{14}$$

If we have a vector space V and consider its n-fold tensor product $\mathcal{T}_n^0(V)$, then S_n act on product states by

$$\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes v_{\sigma_(2)} \otimes \cdots \otimes v_{\sigma(n)}. \tag{15}$$

§. Lie Group and Lie Algebra (General Theory)

Lie group, named after the Norwegian mathematician Sophus Lie, are important in both mathematics and physics because their continuous nature means that we can study group elements that are infinitely close to the identity; these are known to physicists as infinitesimal transformations or generators of the group, and to mathematician as the Lie Algebra. The most general and proper definition requires machinery well outside the scope of this text.

To explore the basic properties of Lie algebra, we define a matrix Lie group to be a subgroup $G \subset GL(n, \mathbb{C})$ which is closed: for any sequence of matrices $A_n \in G$ which converge to a limit matrix A, either $A \in G$ or $A \notin GL(n, \mathbb{C})$. The definition above provide a necessary hypotheses in proving the essential properties of Lie algebras.

Let prove that the groups that we have encountered are indeed matrix Lie group. The orthogonal group $O(n)$ is defined by $AA^T = I$ or $A^T = A^{-1}$. Let's consider the function from $GL(n, \mathbb{C})$ to itself defined by $f(A) = A^T A$. Each entry of matrix $f(A)$ is easily seen to be continuous function of the entries of A, so A is continuous. Consider a new sequence $R_i \subset O(n)$ that converge to R. We then have

$$\begin{aligned} f(A) &= f(\lim R_i) \\ &= \lim f(R_i) \\ &= I. \end{aligned} \tag{16}$$

So $R \in O(n)$, thus $O(n)$ is a Lie group. Lie group can be parameterized in terms of a certain number of real variables. This number is known as the dimension of the Lie group.

For the sake of completeness, we should point out that there are Lie groups out there which are not matrix Lie groups. Their relevance for physics are not established, so we do not consider them here.

Now we would like to zoom into the Lie algebra by consider group elements that are closed to the identity. For concreteness we consider the rotation group $SO(3)$. An arbitrary rotation about the Z axis looks like

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{17}$$

If we take our rotation angle $\epsilon \ll 1$, we can approximate the $R_z(\theta)$ to the first order, which yields

$$R_z(\epsilon) = \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I + \epsilon L_z. \tag{18}$$

where

$$L_Z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{19}$$

Let $\epsilon = \theta/n$. As we take n larger, we expect

$$R_z(\theta) = [R_z(\theta/n)]^n \approx (1 + \frac{\theta L_z}{n})^n. \tag{20}$$

to become an equality in the limit $n \to \infty$. We can write

$$\boxed{R_z(\theta) = e^{\theta L_z}} \tag{21}$$

If we consider $R_z(\theta)$ to be the image of θ under the map

$$R_z: \ \mathbb{R} \to SO(3). \tag{22}$$

Then $R_z(\theta)$ is a homomorphism(Í)! Any such homomorphisms from additive group $\mathbb{R}$ to a matrix Lie group is known as a one-parameter subgroup such as the set of rotation in $\mathbb{R}^3$ about any particular axis and the set of all boosts in a particular direction.

If we have a matrix X such that $e^{tX} \in G, \forall t \in \mathbb{R}$. Then the map

$$\begin{aligned} \exp : \mathbb{R} &\to G \\ t &\mapsto e^{tX} \end{aligned} \tag{23}$$

is a one parameter subgroup. And X is sometimes said to be a generator. For example, the generater of rotation about the z axis L_z is L_z, which is thought to be the second term of the Taylor series,

$$\begin{aligned} R_z(\theta) &= R_z(0) + \theta \frac{dR_z}{d\theta}(0) + \dots \\ &= I + \theta L_z + \dots \end{aligned} \tag{24}$$

Given a matrix Lie group $G \subset GL(n, \mathbb{C})$, we define its Lie algebra $\mathscr{G}$ to be the set of all matrices $X \in M_n(\mathbb{C})$ such that $e^{tX} \in G, \forall t \in \mathbb{R}$. First let's consider $GL(n, \mathbb{C})$ and its Lie algebra $\mathscr{GL}(n, \mathbb{C})$ If X is any element of $M_n(\mathbb{C})$, then e^{tX} is invertible for all t, and its inverse is merely e^{-tX}. Thus

$$\mathscr{GL}(n, \mathbb{C}) = M_n(\mathbb{C}). \tag{25}$$

Second, consider our friend $O(3)$, whose Lie algebra is denoted $\mathscr{O}(3)$. For any $X \in \mathscr{O}(3)$, we have $e^{tX} \in \mathscr{O}(3), \forall t \in \mathbb{R}$, which means

$$(e^{tX})^T e^{tx} = e^{tX^T} e^{tX} = I, \ \forall t. \tag{26}$$

Differentiating this with respect to t yields

$$X^T e^{tX^T} e^{tX} + e^{tX^T} X e^{tX} = 0, \tag{27}$$

which we can evaluate at $t = 0$ to give

$$X^T + X = 0. \tag{28}$$

Thus, any $X \in \mathscr{O}(3)$ must be antisymmetric(we can test it!) and $\mathscr{O}(3)$ is a set of all real antisymmetric matrices. Now $\mathscr{GL}(n, \mathbb{C}), \mathscr{O}(3)$ share a few nice properties.First off, they are all real vector spaces. Secondly, they are closed under commutators, in the sense that if X and Y are elements of the Lie algebra, so is

$$[X, \ Y] = XY - YX. \tag{29}$$

Thirdly, these Lie algebras (all sets of matrices) satisfy the Jacobi identify

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0, \ \forall X, Y, Z \in \mathscr{G} \tag{30}$$

Proposition: Let $\mathscr{G}$ be the Lie Algebra of a matrix Lie group G. Then $\mathscr{G}$ is a real vector space, is closed under commutators, and all elements of $\mathscr{G}$ obey the Jacobi identify.

Proof: Let $\mathscr{G}$ be the Lie algebra of the matrix Lie group and let $X, Y \in \mathscr{G}$. Since $\mathscr{G}$ is closed under the real scalar multiplication, proving that $X + Y \in \mathscr{G}$ will be enough to show that $\mathscr{G}$ is a real vector space. The proof of this hinges on the following identify known as the Lie Product Formula, we state but not prove:

$$e^{X+Y} = \lim_{m\to\infty} (e^{\frac{X}{m}} e^{\frac{Y}{m}})^m. \tag{31}$$

We note that the sequence $A_m = (e^{\frac{X}{m}} e^{\frac{Y}{m}})^m$ is convergent and every element of the sequence is in G. Furthermore, the limit matrix $A = e^{X+Y} \in GL(n, \mathbb{C})$. By the definition of matrix Lie group, then, $e^{X+Y} \in G$.

The second task is to show that $\mathscr{G}$ is closed under the commutators. First we claim that for any $X \in \mathscr{GL}(n, \mathbb{C})$ and $A \in GL(n, .\mathbb{C})$

$$e^{AXA^{-1}} = Ae^X A^{-1}. \tag{32}$$

We can easily verify this by simply expanding the power series on both sides. This implies that if $X \in \mathscr{G}$, then $AXA^{-1} \in \mathscr{G}$ as well. Since $e^{tAXA^{-1}} = Ae^{tX}A^{-1} \in G, \forall t \in \mathbb{R}$. Now let $A = e^{tY}, Y \in \mathscr{G}$, then $e^{tY}Xe^{-tY} \in G, \forall t \in \mathbb{R}$. Noting $\mathscr{G}$ is a vector space, we can calculate the derivative of this equality at $t = 0$

$$\frac{d}{dt}e^{tY}Xe^{-tY}|_{t=0} = YX - XY = \lim_{h\to 0}\frac{e^{hY}Xe^{-hY} - x}{h} \in \mathscr{G}. \tag{33}$$

This shows that the $[Y, X] \in \mathscr{G}$. The verification of Jacobi identity can be calculated directly. Thus the proposition is established.

Before moving forward to the specific Lie algebra, we introduce the Baker-Camphell-Hausdorff(BCH) formula, for which X, Y are sufficient small, e^{X+Y} can be expressed in a single exponential form

$$e^{X+Y} = e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots} \tag{34}$$

The BCH formula demonstrates that the structure of G encodes in the structure of commutators on $\mathscr{G}$.

§. Specific Lie Algebras

Proposition: For any finite-dimensional matrix A we have $\det e^A = e^{\mathrm{Tr}A}$

$\mathscr{O}(n)$ and $\mathscr{U}(n)$: The Lie Algebra of $O(n)$ and $U(n)$

In general, the $\mathscr{O}(n)$ is a set of real, antisymmetric $n \times n$ matrices, so that the dimension is $\dim \mathscr{O}(\mathrm{n}) = \mathrm{n}(\mathrm{n}-1)/2$. As for $U(n)$, $A^\dagger A = I$, and perform the manipulations that led (26) to (28), we have

$$A^\dagger = -A, \forall A \in \mathscr{U}(n) \tag{35}$$

In other word A must be anti-Hermitian. The dimension is n^2 and a real antisymmetric matrix can be considered as an antisymmetric Hermitian matrix, so that $\mathscr{O}(n) \subset \mathscr{U}(n)$.

$\mathscr{O}(n-1, 1)$: The Lie Algebra of $O(n-1, 1)$

Let $X \in \mathscr{O}(n-1, 1)$ be the Lie algebra of the Lorentz group $O(n-1, 1)$. By the definition of Lie algebra and letting $[\eta] = \mathrm{Diag}(1, 1, \dots, 1, -1)$, we have

$$e^{tX^T}[\eta]e^{tX} = [\eta], \forall t \in \mathbb{R} \tag{36}$$

Differentiating with respect to t and evaluate at $t = 0$ we get

$$X^T[\eta] + [\eta]X = 0 \tag{37}$$

Writing X in block form with an $(n-1) \times (n-1)$ matrix X' for the spatial components($X_{ij}, i, j < n$) and vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ for the components of $X_{i,n}$ and $X_{n,i}$ (i<n). The above equation reads

$$\begin{pmatrix} X'^T & -\boldsymbol{b} \\ \boldsymbol{a} & -X_{nn} \end{pmatrix} + \begin{pmatrix} X' & \boldsymbol{a} \\ -\boldsymbol{b} & -X_{nn} \end{pmatrix} = 0 \tag{38}$$

which implies that X has the form

$$X = \begin{pmatrix} X' & \boldsymbol{a} \\ \boldsymbol{a} & 0 \end{pmatrix},\ X' \in \mathscr{O}(n-1), \boldsymbol{a} \in \mathbb{R}^{n-1} \tag{39}$$

We will discuss the situation in four dimensions in the next section.

$s\mathscr{O}(3)$ **and** $s\mathscr{U}(2)$

As discussed above, $s\mathscr{O}(2)$ consists all of the antisymmetric matrices in the form

$$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \tag{40}$$

so $s\mathscr{O}(3)$ is one-dimensional and we may take a basis

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{41}$$

Explicitly we can computer that

$$e^{\theta X} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \theta \in \mathbb{R}. \tag{42}$$

As mentioned above, the basis for $s\mathscr{O}(3)$ can be chosen as L_x, L_y, L_z:

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, L_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{43}$$

Then, we can write the element $X \in s\mathscr{O}(3)$ as

$$X = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} = xL_x + yL_y + zL_z. \tag{44}$$

We can check the commutators of basis elements work out to be

$$\boxed{[L_i, L_j] = \sum_{k=1,2,3} \epsilon_{ijk} L_k.} \tag{45}$$

Now if we take $[X]$ as a unit vector and rename it $\boldsymbol{n}$, we can prove that $e^{\theta X}$ indeed represents the rotation about $\boldsymbol{n}$ by an angle θ.

$s\mathscr{U}(2)$ is a set of all 2×2 traceless anti-Hermitian matrices. We take the normalized Pauli matrices as the basis

$$S_x = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, S_y = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, S_z = \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}. \tag{46}$$

Also, we can check the commutator relations

$$\boxed{[S_i, S_j] = \sum_{k=1,2,3} \epsilon_{ijk} S_k.} \tag{47}$$

From direct calculation, we can show that

$$\boxed{e^{\boldsymbol{\theta}\cdot\boldsymbol{S}} = \cos(\theta/2)I + 2\boldsymbol{S}\cdot\boldsymbol{n}\sin(\theta/2)} \tag{48}$$

where $\boldsymbol{\theta} = \theta\boldsymbol{n}$ and $\boldsymbol{n}$ is a unit vector. In reality,$e^{\boldsymbol{\theta}\cdot\boldsymbol{S}}$ represents the rotation about $\boldsymbol{n}$ by the angle θ, which means $U = e^{\boldsymbol{\theta}\cdot\boldsymbol{S}} \leftrightarrow R = e^{-\boldsymbol{\theta}\cdot\boldsymbol{L}}$ and we will prove this later.

$s\mathscr{O}(3,1)$
We know that an arbitrary element in $s\mathscr{O}(3,1)$ can be written in the form

$$X = \begin{pmatrix} X' & \boldsymbol{a} \\ -\boldsymbol{a} & 0 \end{pmatrix} \tag{49}$$

with $X' \in s\mathscr{O}(3)$ and $a \in \mathbb{R}^3$. Embedding the $\boldsymbol{L}$ of $s\mathscr{O}(3)$ into $s\mathscr{O}(3,1)$ as

$$\tilde{L}_i = \begin{pmatrix} L_i & \boldsymbol{0} \\ \boldsymbol{0} & 0 \end{pmatrix} \tag{50}$$

and defining new generators K_i:

$$K_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{51}$$

we have the following commutation equations:

$$\begin{aligned} [\tilde{L}_i, \tilde{L}_j] &= \sum_{k=1,2,3} \epsilon_{ijk}\tilde{L}_k \\ [\tilde{L}_i, K_j] &= \sum_{k=1,2,3} \epsilon_{ijk}K_k \\ [K_i, K_j] &= -\sum_{k=1,2,3} \epsilon_{ijk}K_k. \end{aligned} \tag{52}$$

In another aspect, K_i can be interpreted as generating boosts along their corresponding axes and we can show that $e^{u^i K_i}$ represents a boost in the direction of $\boldsymbol{u}$.

§. The Dirac Equation

Rotation and boost

We now construct a traceless 2×2 matrix transforming under $SU(2)$:

$$h = \boldsymbol{\sigma}\cdot\boldsymbol{r} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}. \tag{53}$$

where the matrices $\boldsymbol{\sigma}$ are well-known Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \tag{54}$$

h is a Hermitian and the transformation

$$h \to UhU^{\dagger} = h^{'} = \begin{pmatrix} z^{'} & x^{'} - iy^{'} \\ x^{'} + iy^{'} & -z^{'} \end{pmatrix}. \tag{55}$$

By solving the transformation between $(x, y, z)^T$ and $(x^{'}, y^{'}, z^{'})^T$, we can find that the $SU(2)$ is isomorphic to $SO(3)$, in another word

$$U = e^{i\boldsymbol{\theta}\cdot\boldsymbol{\sigma}/2} \leftrightarrow R = e^{i\boldsymbol{\theta}\cdot\boldsymbol{J}} \tag{56}$$

where $\boldsymbol{J} = i\boldsymbol{L}$.

We can define an action of $SL(2, \mathbb{C})$ on Minkowski spacetime by writing a point of space time as a two-by-two Hermitian matrix in the form

$$X = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix}. \tag{57}$$

This presentation has the feature that $\det X = t^2 - x^2 - y^2 - z^2$ which is supposed to be invariant under the Lorentz transformation. $SL(2, \mathbb{C})$ acts on the space of Hermitian matrices via

$$X \mapsto PXP^{\dagger} \tag{58}$$

and this action preserves the determinant. Now, we define a transformation matrix $[S]$, which satisfies

$$X^{'} = (t^{'}, x^{'}, y^{'}, z^{'})^T = SX = S(t, x, y, z)^T. \tag{59}$$

Under these definitions, the relations between $SL(2, \mathbb{C})$ and boost and rotation are established

$$\boxed{\begin{aligned} &e^{i\boldsymbol{\theta}\cdot\boldsymbol{\sigma}/2} \leftrightarrow e^{-\boldsymbol{\theta}\cdot\tilde{L}} \text{ (Rotation)} && (60) \\ &e^{\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} \leftrightarrow e^{\boldsymbol{\phi}\cdot\boldsymbol{K}} \text{ (Pure Lorentz transformation)} && (61) \end{aligned}}$$

The jointed effect of boost and rotation on a spinor ξ gives

$$\xi \to \exp\left[\frac{i\boldsymbol{\sigma}}{2} \cdot (\boldsymbol{\theta} - \boldsymbol{i\phi})\right] \xi \tag{62}$$

and this transformation is the first step to construct the Dirac equation.

Let's do this!

At first, define the generators:

$$\begin{aligned} \boldsymbol{A} &= \frac{1}{2}(\tilde{\boldsymbol{L}} + i\boldsymbol{K}) \\ \boldsymbol{B} &= \frac{1}{2}(\tilde{\boldsymbol{L}} - i\boldsymbol{K}) \end{aligned} \tag{63}$$

We denote the spinor ξ. If $(\boldsymbol{\theta}, \boldsymbol{\phi})$ are parameters of rotation and pure Lorentz transformation, ξ transforms under Type I in equation (63) as

$$\xi \to \exp\left[\frac{i\boldsymbol{\sigma}}{2} \cdot (\boldsymbol{\theta} - \boldsymbol{i\phi})\right] \xi. \tag{64}$$

Under Type II, the spinor η transforms as

$$\eta \to \exp\left[\frac{i\boldsymbol{\sigma}}{2} \cdot (\boldsymbol{\theta} + \boldsymbol{i\phi})\right] \eta. \tag{65}$$

If we consider the parity, it is not sufficient to consider ξ and η separately, but the four-spinor

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix}. \tag{66}$$

Under Lorentz transformation ψ transforms as

$$\boxed{\begin{pmatrix} \xi \\ \eta \end{pmatrix} \to \begin{pmatrix} e^{1/2\boldsymbol{\sigma}\cdot(\boldsymbol{\theta}-\boldsymbol{i\phi})} & 0 \\ 0 & e^{1/2\boldsymbol{\sigma}\cdot(\boldsymbol{\theta}+\boldsymbol{i\phi})} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}} \tag{67}$$

Now let us specialize the equation above to the case of Lorentz boost($\boldsymbol{\theta} = 0$), and at the same time relabel the two-spinor ξ, η to ϕ_R, ϕ_L respectively. R and L standing for their chirality. We have

$$\begin{aligned} \phi_R &\to e^{1/2\boldsymbol{\sigma}\cdot\boldsymbol{\phi}} \phi_R \\ &= [\cosh(\phi/2) + \boldsymbol{\sigma} \cdot \boldsymbol{n} \sinh(\phi/2)]\phi_R \end{aligned} \tag{68}$$

where $\boldsymbol{n}$ is the unit vector in the direction of Lorentz boost. Suppose that the initial spinor refers to a particle at rest, $\phi_R(0)$, and the transformed one to a particle with momentum $\boldsymbol{p}$, $\phi_R(\boldsymbol{p})$. From the special relativity, we have $\cosh(\phi/2) = \sqrt{(\gamma+1)/2}$, $\sinh(\phi/2) = \sqrt{(\gamma-1)/2}$, so

$$\begin{aligned} \phi_R(\boldsymbol{p}) &= \left[(\frac{\gamma+1}{2})^{1/2} + \boldsymbol{\sigma} \cdot \hat{\boldsymbol{p}}(\frac{\gamma-1}{2})^{1/2}\right] \phi_R(0) \\ &= \frac{E + m + \boldsymbol{\sigma} \cdot \boldsymbol{p}}{[2m(E+m)]^{1/2}} \phi_R(0) \end{aligned} \tag{69}$$

noting $\gamma = E/m$ ($c = 1$). When the particle is at rest we cannot identify it to be right or left, so $\phi_R(0) = \phi_L(0)$. Then it follows, after some algebra

$$\begin{pmatrix} -m & p_0 + \boldsymbol{\sigma} \cdot \boldsymbol{p} \\ p_0 - \boldsymbol{\sigma} \cdot \boldsymbol{p} & -m \end{pmatrix} \begin{pmatrix} \phi_R(\boldsymbol{p}) \\ \phi_L(\boldsymbol{p}) \end{pmatrix} = 0 \tag{70}$$

Defining the 4×4 matrices

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}, \tag{71}$$

equation (70) becomes

$$\boxed{(\gamma^u p_u - m)\psi(p) = 0.} \tag{72}$$

This is the Dirac equation (in natural unit $c = 1, \hbar = 1$) for massive spin-$\frac{1}{2}$ particles! Hia hia hia...

Then we have antiparticles, correct electron magnetic moment, and spin-orbit interactions with the correct Thomas precession factor of 2 etc.

References

Group theory and Lie group:

An introduction to tensors and group theory for physicists, Nadir Jeevanjee.

Boost and rotation, Lorentz group, the Dirac equation:

Quantum Field Theory(second edition), Lewis H. Ryder.

Wikipedia: https://en.wikipedia.org/wiki/Lorentz_group