SU(2) to SO(3) homomorphism

Nguyen Duc Thanh, Yesui Baatar

(Fall 2022)

1 Introduction

In this report, we want to show that $\mathrm{SU}(2)/C_2 \simeq \mathrm{SO}(3)$. This result follows from finding a surjective homomorphism

$$\phi : \quad \mathrm{SU}(2) \to \mathrm{SO}(3),$$

with $\mathrm{Ker}(\phi) = \{\mathbb{I}, -\mathbb{I}\} \simeq C_2$ where $\mathbb{I}$ is the identity element of SU(2). In addition, we will show that one choice of such homomorphism is given by

$$[\phi(U)]_{ij} = \frac{1}{2}\mathrm{Tr}(\sigma_i U \sigma_j U^{-1}),$$

where σ_i's are the Pauli matrices.

2 Derivation

Recall the Pauli matrices with the identity

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

They form a basis of $M_2(\mathbb{C})$ since they are linearly independent and span $M_2(\mathbb{C})$. Let us consider the subspace $V \subset M_2(\mathbb{C})$ spanned by $\sigma_1, \sigma_2, \sigma_3$. This means that any element $A \in V$ can be written as

$$A = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3 = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}.$$

If x_i's are real, we can consider A as a map

$$\begin{aligned} A : \quad & \mathbb{R}^3 \to V, \\ & \mathbf{x} \mapsto x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3. \end{aligned}$$

One can check that this map is bijective. Then, for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, we have

$$A(\mathbf{x})A(\mathbf{y}) = (x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)(y_1\sigma_1 + y_2\sigma_2 + y_3\sigma_3) = \sum_{i,j} x_i y_j \sigma_i \sigma_j,$$

where $i, j = 1, 2, 3$. Upon taking the trace, we get

$$\mathrm{Tr}[A(\mathbf{x})A(\mathbf{y})] = \sum_{i,j} x_i y_j \mathrm{Tr}(\sigma_i \sigma_j).$$

The Pauli matrices have the following properties

$$\sigma_i^2 = \mathbb{I} \quad \forall i,$$

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 0 \quad \forall i \neq j.$$

By taking the trace of these equations and using the cyclic property of the trace, it can be shown that

$$\mathrm{Tr}(\sigma_i^2) = 2 \quad \forall i, \quad \text{and} \quad \mathrm{Tr}(\sigma_i \sigma_j) = 0 \quad \forall i \neq j,$$

$$\text{or equivalently} \quad \mathrm{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}.$$

Then, we get

$$\mathrm{Tr}[A(\mathbf{x})A(\mathbf{y})] = 2\sum_i x_i y_i = 2\langle \mathbf{x}, \mathbf{y} \rangle,$$

where $\langle \, , \rangle$ is the inner product defined in $\mathbb{R}^3$. Therefore, we can define an "inner product" $\langle \, , \rangle$ in V that is equivalent to the one in $\mathbb{R}^3$ by

$$\langle A(\mathbf{x}), A(\mathbf{y}) \rangle = \frac{1}{2} \mathrm{Tr}[A(\mathbf{x})A(\mathbf{y})].$$

Now, let us define a transformation T_U on V where $U \in \mathrm{SU}(2)$ by

$$T_U[A(\mathbf{x})] = UA(\mathbf{x})U^{-1} = UA(\mathbf{x})U^*.$$

The map T_U, indeed, gives us an element of V and we will show it as follows. Consider V as a subspace of $M_2(\mathbb{C})$ in which every element is self-adjoint and has zero trace. These properties completely define V the same as before. Then, one has

$$\begin{cases} (T_U[A(\mathbf{x})])^* = [UA(\mathbf{x})U^*]^* = UA(\mathbf{x})U^* = T_U[A(\mathbf{x})], \\ \mathrm{Tr}\,(T_U[A(\mathbf{x})]) = \mathrm{Tr}[UA(\mathbf{x})U^{-1}] = \mathrm{Tr}[U^{-1}UA(\mathbf{x})] = \mathrm{Tr}[A(\mathbf{x})] = 0, \end{cases}$$

so $T_U[A(\mathbf{x})] \in V$. We can now consider the inner product of $T_U[A(\mathbf{x})]$ and $T_U[A(\mathbf{y})]$,

$$\begin{aligned} \langle T_U[A(\mathbf{x})], T_U[A(\mathbf{y})] \rangle &= \frac{1}{2}\mathrm{Tr}[UA(\mathbf{x})U^{-1}UA(\mathbf{y})U^{-1}] \\ &= \frac{1}{2}\mathrm{Tr}[UA(\mathbf{x})A(\mathbf{y})U^{-1}] \\ &= \frac{1}{2}\mathrm{Tr}[U^{-1}UA(\mathbf{x})A(\mathbf{y})] \\ &= \frac{1}{2}\mathrm{Tr}[A(\mathbf{x})A(\mathbf{y})] \\ &= \langle A(\mathbf{x}), A(\mathbf{y}) \rangle. \end{aligned}$$

This means that T_U preserves the inner product in V. However, consider the transformation on $\mathbb{R}^3$ instead of V. Set $\phi_U = A^{-1} \circ T_U \circ A$ which is a transformation on $\mathbb{R}^3$ corresponding to T_U. Then, the above equality implies

$$\langle \phi_U(\mathbf{x}), \phi_U(\mathbf{x}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle,$$

in $\mathbb{R}^3$ so ϕ_U preserves the inner product in this space. As a result, $\phi_U \in \mathrm{O}(3)$. The relations between the maps and the spaces can be summarized by the following diagram

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{\phi_U} & \mathbb{R}^3 \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle A} \\ V & \xrightarrow[T_U]{} & V \end{array}$$

We can also consider U as a variable and ϕ as a (continuous) map, so we have

$$\begin{aligned} \phi: \quad & \mathrm{SU}(2) \to \mathrm{O}(3), \\ & U \mapsto \phi_U. \end{aligned}$$

Proving that ϕ is a group homomorphism:

To show that ϕ is a homomorphism with the properties as mentioned in Section 1, we need to show that ϕ preserves the group law and the identity in $\mathrm{SU}(2)$ is mapped to the identity in $\mathrm{SO}(3)$. By definition,

$$\phi_U(\mathbf{x}) = A^{-1}(T_U[A(\mathbf{x})]) = A^{-1}[UA(\mathbf{x})U^*] \quad \text{or} \quad UA(\mathbf{x})U^* = A[\phi_U(\mathbf{x})].$$

Then, for $U, V \in \mathrm{SU}(2)$,

$$\begin{aligned} \phi_{UV}(\mathbf{x}) &= A^{-1}[(UV)A(\mathbf{x})(UV)^*] \\ &= A^{-1}[UVA(\mathbf{x})V^*U^*] \\ &= A^{-1}(UA[\phi_V(\mathbf{x})]U^*) \\ &= \phi_U(\phi_V(\mathbf{x})), \end{aligned}$$

meaning that $\phi_{UV} = \phi_U \phi_V$. Also, we have $\mathbb{I}A(\mathbf{x})\mathbb{I}^* = A(\mathbf{x})$ so

$$\phi_{\mathbb{I}}(\mathbf{x}) = A^{-1}A(\mathbf{x}) = \mathbf{x},$$

implying that $\phi_{\mathbb{I}} = \mathbb{I}'$, in which we have denoted $\mathbb{I}'$ as the identity element in $\mathrm{O}(3)$. These prove that ϕ is a homomorphism.

Finding the range of ϕ:

Let us return to the statement that $\phi_U \in \mathrm{O}(3)$. This does not necessarily mean that every element of $\mathrm{O}(3)$ can be reached by ϕ, so we want to restrict the codomain of ϕ to be $\mathrm{Ran}(\phi)$ only. First, we notice that every element of $\mathrm{SU}(2)$ can be written as

$$U = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix},$$

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. Since we can identify any element $z \in \mathbb{C}$ by an element $(z_1, z_2) = (\mathrm{Re}(z), \mathrm{Im}(z)) \in \mathbb{R}^2$, we have

$$|\alpha|^2 + |\beta|^2 = \alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2 = 1,$$

which corresponds to 3-dimensional spherical surface S^3 in $\mathbb{R}^4$. Therefore, SU(2) is homeomorphic to S^3, and every continuous path in SU(2) can be identified with a continuous path on this surface. Because S^3 is connected, we also find SU(2) to be connected.

Now, we will prove that the range of a continuous map from a connected space is also connected. Let f be a continuous and surjective map

$$f : X \to Y,$$

where X is connected. Assume that Y is not connected, i.e., there exists A, B non-empty and open in Y such that $A \cup B = Y$ and $A \cap B = \varnothing$. From the definitions, we have

$$\begin{cases} A, B \text{ are non-empty} \Rightarrow f^{-1}(A) \text{ and } f^{-1}(B) \text{ are non-empty,} \\ f \text{ is a function} \Rightarrow f^{-1}(A) \cap f^{-1}(B) = \varnothing, \\ f \text{ is continuous} \Rightarrow f^{-1}(A) \text{ and } f^{-1}(B) \text{ are open,} \\ f \text{ is surjective} \Rightarrow f^{-1}(A) \cup f^{-1}(B) = X. \end{cases}$$

This implies that X is not connected as we can separate X into disjoint non-empty open subsets $f^{-1}(A)$ and $f^{-1}(B)$, leading to a contradiction. Thus, $Y = \mathrm{Ran}(f)$ is connected. For our problem, this means that $\mathrm{Ran}(\phi)$ is connected. Since $\phi_{\mathbb{I}} = \mathbb{I}'$ and ϕ preserves the group law, $\mathrm{Ran}(\phi)$ is a subgroup of SO(3) — the identity component of O(3) — because it is connected, in contrary to O(3). The outline for the proof on the connectedness of SO(3) (or general SO(n)) is given in Chapter 1, Exercise 13 of [3], whereas and the non-connectedness of O(3) can be seen from the fact that we can make the separation

$$\mathrm{O}(3) = \mathrm{SO}(3) \cup I \cdot \mathrm{SO}(3), \text{ where } I = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Because $\det(\mathrm{SO}(3)) = 1$ while $\det(I \cdot \mathrm{SO}(3)) = -1$, we cannot define a continuous path joining the two subsets so they are disjoint, implying that O(3) is not connected.

We have shown that $\mathrm{Ran}(\phi) \subseteq \mathrm{SO}(3)$; in fact, it can also be shown that $\mathrm{SO}(3) \subseteq \mathrm{Ran}(\phi)$. The proof of the latter statement is given in the Appendix at the end of the report. As a result, we get $\mathrm{Ran}(\phi) = \mathrm{SO}(3)$ so the map

$$\phi : \quad \mathrm{SU}(2) \to \mathrm{SO}(3)$$

is surjective.

Finding the kernel of ϕ:

By definition, if $U_0 \in \mathrm{Ker}(\phi)$, we have $\phi_{U_0} = \mathbb{I}'$. This means that

$$\phi_{U_0}(\mathbf{x}) = \mathbf{x} \quad \text{or} \quad U_0 A(\mathbf{x}) U_0^* = A(\mathbf{x})$$

for arbitrary $\mathbf{x} \in \mathbb{R}^3$. Since $A(\mathbf{x}) = \sum x_i \sigma_i$, the above equation is equivalent to

$$U_0 \sigma_i U_0^* = \sigma_i \quad \forall i.$$

As mentioned previously, we can write U_0 as

$$U_0 = \begin{pmatrix} \alpha_0 & -\bar{\beta}_0 \\ \beta_0 & \bar{\alpha}_0 \end{pmatrix},$$

with $|\alpha_0|^2 + |\beta_0|^2 = 1$. Let us directly calculate each case

$$U_0 \sigma_1 U_0^* = \sigma_1 \iff \begin{pmatrix} -\alpha_0\beta_0 - \bar{\alpha}_0\bar{\beta}_0 & \alpha_0^2 - \bar{\beta}_0^2 \\ \bar{\alpha}_0^2 - \beta_0^2 & \alpha_0\beta_0 + \bar{\alpha}_0\bar{\beta}_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$U_0 \sigma_2 U_0^* = \sigma_2 \iff i \begin{pmatrix} \alpha_0\beta_0 - \bar{\alpha}_0\bar{\beta}_0 & -\alpha_0^2 - \bar{\beta}_0^2 \\ \bar{\alpha}_0^2 + \beta_0^2 & -\alpha_0\beta_0 + \bar{\alpha}_0\bar{\beta}_0 \end{pmatrix} = i \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$U_0 \sigma_3 U_0^* = \sigma_3 \iff \begin{pmatrix} |\alpha_0|^2 - |\beta_0|^2 & 2\alpha_0\bar{\beta}_0 \\ 2\bar{\alpha}_0\beta_0 & -|\alpha_0|^2 + |\beta_0|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

From the off-diagonal elements in the third equation, we find that either $\alpha_0 = 0$ or $\beta_0 = 0$. However, if $\alpha_0 = 0$, we have $|\beta_0| = 1$ so the third equation becomes

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which is not true. Thus, we are left with $\beta_0 = 0$ so $|\alpha_0| = 1$ and the set of equations reads

$$\begin{cases} \begin{pmatrix} 0 & \alpha_0^2 \\ \bar{\alpha}_0^2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & -\alpha_0^2 \\ \bar{\alpha}_0^2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{(always true)} \end{cases}.$$

Solving these equations give us

$$\alpha_0^2 = \bar{\alpha}_0^2 = 1 \iff \alpha_0 \in \{1, -1\}.$$

As a result, we have $\mathrm{Ker}(\phi) = \{\mathbb{I}, -\mathbb{I}\} \simeq C_2$.

Finding the explicit expression of ϕ:

Consider the equation

$$UA(\mathbf{x})U^* = A[\phi_U(\mathbf{x})].$$

We can expand the left-hand side as

$$UA(\mathbf{x})U^* = \sum_{i=1}^{3} x_i U \sigma_i U^*,$$

while the right-hand side is given by

$$A[\phi_U(\mathbf{x})] = \sum_{i=1}^{3} [\phi_U(\mathbf{x})]_i \sigma_i = \sum_{i=1}^{3} \sum_{j=1}^{3} [\phi_U]_{ij} x_j \sigma_i.$$

We take the summation in the left-hand side equation to be over j since it is just a dummy index. Then, by collecting terms with the same j, one has

$$\sum_{j=1}^{3} x_j \left(\sum_{i=1}^{3} [\phi_U]_{ij} \sigma_i - U \sigma_j U^* \right) = 0.$$

This has to hold for any choice of x_j's so in general, we have

$$\sum_{i=1}^{3} [\phi_U]_{ij} \sigma_i - U \sigma_j U^* = 0 \quad \text{or} \quad \sum_{i=1}^{3} [\phi_U]_{ij} \sigma_i = U \sigma_j U^*.$$

Now, let σ_k for some k act on the right of the second equation then take the trace

$$\sum_{i=1}^{3} [\phi_U]_{ij} \mathrm{Tr}(\sigma_i \sigma_k) = \mathrm{Tr}(U \sigma_j U^* \sigma_k).$$

By using $\mathrm{Tr}(\sigma_i \sigma_k) = 2\delta_{ik}$, we get

$$2[\phi_U]_{kj} = \mathrm{Tr}(U \sigma_j U^* \sigma_k) \quad \text{or} \quad [\phi_U]_{ij} = \frac{1}{2} \mathrm{Tr}(\sigma_i U \sigma_j U^*),$$

after cycling to the right inside the trace once and changing the index from k to i. This is the same expression given in Section 1 since $U^{-1} = U^*$. Another short way to derive this is by directly calculating the matrix element using its definition with the inner product and the unit vectors $\mathbf{e}_j$'s in $\mathbb{R}^3$, i.e.,

$$\begin{aligned}
[\phi_U]_{ij} &= \langle \mathbf{e}_i, \phi_U(\mathbf{e}_j) \rangle && (\text{in } \mathbb{R}^3) \\
&= \langle A(\mathbf{e}_i), A[\phi_U(\mathbf{e}_j)] \rangle && (\text{in } V) \\
&= \langle \sigma_i, U \sigma_j U^* \rangle \\
&= \frac{1}{2} \mathrm{Tr}(\sigma_i U \sigma_j U^*).
\end{aligned}$$

□

Let us make a final remark on this result. The homeomorphism of SU(2) and S^3 in $\mathbb{R}^4$ also implies that SU(2) is simply connected (while SO(3) is not). Then, the above isomorphism means that we can relate problems regarding the non-simply connected group SO(3) with the simply connected group SU(2). We say that SU(2) is a universal cover of SO(3). Simple connectedness is of great importance in Lie groups and Lie algebras because if a Lie group is simply connected, there is a one-to-one correspondence between its representation (or homomorphism) and the representation (or homomorphism) of its Lie algebra (see section 3.6 and 3.7, [3]) .

3 References

1. Groups and their representations, lecture notes by S. Richard.
2. [A] Theorie des groupes pour la physique, lecture notes by W. Amrein.
3. [H] Lie Groups, Lie Algebras, and Representations: An Elementary Introduction, book by B. Hall.
4. Topology, book by J. R. Munkres.
5. Introduction to Lie Groups and Lie Algebras, book by A. A. Sagle and R. E. Walde.

Appendix

We will first work with topological groups for generality then apply it to our problem. Let G be a topological group with e be its identity and G_0 be its identity component. A neighborhood V of e is called symmetric if

$$V^{-1} := \{v^{-1} | v \in V\} = V.$$

Let V be open and define $V^k := VV \cdots V$ (k times). Then, consider the following subset of G

$$H := \bigcup_{k=1}^{\infty} V^k.$$

For any $x \in V^m \subset H$ and $y \in V^n \subset H$, we have $xy \in V^{m+n} \subset H$, and $x^{-1} \in (V^{-1})^m = V^m \subset H$. Also, the conditions for associativity and existence of identity are satisfied because $H \subset G$ which is a group itself and V contains the identity. Thus, H is a subgroup of G. Assume that V^k is open for $k \geq 1$, let us consider

$$V^{k+1} = VV^k = \bigcup_{a \in V} aV^k.$$

which is a union of left cosets of V^k (we can also write in terms of right cosets by considering $V^{k+1} = V^k V$ instead). For arbitrary $x \in G$, we define a left translation of x on G by

$$\begin{aligned} L(x): \quad & G \to G, \\ & y \mapsto xy. \end{aligned}$$

G is a topological group so by definition, the product map and inverse map for the group are continuous. Therefore, both $L(x)$ and $L(x)^{-1} = L(x^{-1})$ are continuous maps [1]. If we let $x = a \in V$, the map $L(a)^{-1}$ is continuous so by definition, the inverse image of an open set in G is also an open set in G. Since $\left(L(a)^{-1}\right)^{-1} = L(a)$ and V^k is open, we have $L(a)V^k = aV^k$ is open. Then, V^{k+1} is open due to it being a union of open sets. We have $V^1 = V$ open so by induction, V^k is open for any k. Then, H is a union of open sets so H is open as well. However, using the argument for the left coset again, we can show that the subset

$$K := \bigcup_{b \notin H} bH = G \setminus H$$

is open, which means that H is closed. As a result, H is an open and closed subgroup of G. In particular, we consider $G = G_0$, which is connected. For a connected space, the only subsets that are both open and closed are the whole set itself and the empty set, and because H is non-empty by assumption, we have $H = G_0$. The readers can refer to §23 of [4] for the proof of this statement.

[1] In fact, the map $L(x)$ is a homeomorphism.

The above derivation only involves symmetric neighborhoods of e but we can make this more general as follows. We consider the identity component G_0 alone and let $U \subseteq G_0$ be any open neighborhood of e. From Lemma 3.4 in [5], we can always find a neighborhood $V \subseteq U$ of e such V is open and symmetric. Then, we have $V^k \subseteq U^k$ and

$$H = \bigcup_{k=1}^{\infty} V^k \subseteq \bigcup_{k=1}^{\infty} U^k.$$

As shown before, $H = G_0$, and because $U^k \subseteq G_0$ for any k, we arrive at

$$G_0 = \bigcup_{k=1}^{\infty} U^k.$$

Thus, if we have a connected topological group G, i.e., $G = G_0$, any open neighborhood U of e is a generator of G.

Now, let us return to our problem. We will consider the map

$$\phi : \quad \mathrm{SU}(2) \to \mathrm{SO}(3).$$

We state without proof that there exists (small) open neighborhoods $A \subseteq \mathrm{SU}(2)$ of $\mathbb{I}$ and $B \subseteq \mathrm{SO}(3)$ of $\mathbb{I}'$ such that ϕ defined on these subsets is a homeomorphism. This also means that $B \subseteq \mathrm{Ran}(\phi)$, and since ϕ is also a group homomorphism, we have $B^k \subseteq \mathrm{Ran}(\phi)$ for any k. Then, taking the union over all k gives us

$$\bigcup_{k=1}^{\infty} B^k \subseteq \mathrm{Ran}(\phi).$$

From the result derived previously, the left-hand side is equal to SO(3) itself, so finally, we have $\mathrm{SO}(3) \subseteq \mathrm{Ran}(\phi)$.