QSENN: Quantum Spin Equivariant Neural Network

Abhijatmedhi Chotrattanapituk

Quantum Measurement Group, MIT, Cambridge, MA, USA

Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, USA

(Dated: May 26, 2023)

1. INTRODUCTION

2. NOTATIONS

Due to the number of groups, and representations discussed in this work, it is important to define some notations that would be used.

3. $SU(2)$'S DEFINITION

$SU(2)$ is the special unitary group of dimension 2, i.e. the group of 2 by 2 unitary matrices with unit determinant. In other words,

$$SU(2) = \{u \in GL(2,\mathbb{C})|u^\dagger = u^{-1}, \det(u) = 1\}, \tag{1}$$

where $GL(2,\mathbb{C})$ is the group of 2 by 2 invertible matrices with complex field, and $u^\dagger$ is the complex conjugate transposition of u. Since for any $u, v \in SU(2)$

$$\det(uv) = \det(u)\det(v) = 1, \tag{2}$$

and

$$(uv)^\dagger = (uv)^{\top*} = (v^\top u^\top)^* = v^\dagger u^\dagger = v^{-1}u^{-1} = (uv)^{-1}, \tag{3}$$

$SU(2)$ is closed subgroup of $GL(2,\mathbb{C})$ which makes it a Lie group. This means that there is a corresponding Lie algebra, $\mathfrak{su}(2)$, such that

$$SU(2) = \{e^g|g \in \mathfrak{su}(2)\}. \tag{4}$$

From definition of matrix exponential,

$$(e^g)^\dagger = e^{g^\dagger},$$

$$(e^g)^{-1} = e^{-g}, \tag{5}$$

$$\det(e^g) = e^{\operatorname{Tr} g},$$

one can define the Lie algebra of $SU(2)$ as

$$\mathfrak{su}(2) = \{g \in M(2, \mathbb{C}) | g^\dagger = -g, \operatorname{Tr}(g) = 0\}, \tag{6}$$

where $M(2, \mathbb{C})$ is the group of 2 by 2 matrices with complex field, and $\operatorname{Tr}(g)$ is the trace of matrix g. Hence, each member g of $\mathfrak{su}(2)$ can be written in the From

$$g = \begin{bmatrix} iv_z & v_y + iv_x \\ -v_y + iv_x & -iv_z \end{bmatrix} = iv_x\sigma_x + iv_y\sigma_y + iv_z\sigma_z = i\vec{v} \cdot \vec{\sigma} \tag{7}$$

where $\vec{v} \in \mathbb{R}^3$, and σ_i's are Pauli matrices.

4. $\mathfrak{su}(2)$'S IRREDUCIBLE REPRESENTATION

From $\mathfrak{su}(2)$, one can consider Pauli matrices as the group generators, but it is more common to use $J_i^{(1/2)} = \sigma_i/2$ as the generators. This gives the Lie bracket relation as

$$\left[J_i^{(1/2)}, J_j^{(1/2)} \right] = i\epsilon_{ijk} J_k^{(1/2)}. \tag{8}$$

The next step is to consider an arbitary representation of $SU(2)$ such that its generators, J_i's, obey (8). With out loss of generality, consider the representation vector space with eigenvectors of J_z, $\{|\lambda_{J_z}\rangle\}$, as the basis,

$$J_z |\lambda_{J_z}\rangle = \lambda_{J_z} |\lambda_{J_z}\rangle. \tag{9}$$

Furthermore, we will also consider only the representation that is irreducible. With the standard method, first define two additional operators

$$J_+ = J_x + iJ_y,$$

$$j_- = J_x - iJ_y. \tag{10}$$

From Lie bracket, and the choice of basis used,

$$J_z J_+ |\lambda_{J_z}\rangle = J_+(J_Z + 1)|\lambda_{J_z}\rangle = (\lambda_{J_z} + 1)J_+ |\lambda_{J_z}\rangle,$$

$$J_z J_- |\lambda_{J_z}\rangle = J_-(J_Z - 1)|\lambda_{J_z}\rangle = (\lambda_{J_z} - 1)J_- |\lambda_{J_z}\rangle. \tag{11}$$

This directly implies

$$J_+ \ket{\lambda_{J_z}} \propto \ket{\lambda_{J_z} + 1},$$
$$J_- \ket{\lambda_{J_z}} \propto \ket{\lambda_{J_z} - 1}. \tag{12}$$

Before proceeding, we need to clarify an assumption we made to get (12) that there is no degeneracy of λ_{J_z} in the basis. This result is directly entailed from the assumption that the representation is irreducuble. To prove this, first assume that there exist $\ket{\lambda_{J_z}}'$ that has the same eigenvalue as $\ket{\lambda_{J_z}}$. From the properties of J_i's, we can choose the representation such that (12) is true. This means that the proper subspace that exclude $\ket{\lambda_{J_z}}'$ is invariant for this choice of representation. Hence, the representation is reducible which contradict the assumption.

To get the proportionality, consider another operators

$$J^2 = J_x^2 + J_y^2 + J_z^2. \tag{13}$$

One can directly chack that the new operator commutes with all J_i's mentioned so far. Furthermore, it is straight forward that

$$\begin{aligned} J^2 &= J_+ J_- + J_z^2 - J_z \\ &= J_- J_+ + J_z^2 + J_z \end{aligned} \tag{14}$$

Since J^2 commutes with J_z, the basis used are also eigenbasis of J^2 with eigenvalue λ_{J^2}. Hence, we can directly relabel the eigenbasis from $\ket{\lambda_{J_z}}$ to $\ket{\lambda_{J^2}, \lambda_{J_z}}$, and get

$$\bra{\lambda_{J^2}, \lambda_{J_z}} J_+ J_- \ket{\lambda_{J^2}, \lambda_{J_z}} = \lambda_{J^2} - \lambda_{J_z}(\lambda_{J_z} - 1),$$
$$\bra{\lambda_{J^2}, \lambda_{J_z}} J_- J_+ \ket{\lambda_{J^2}, \lambda_{J_z}} = \lambda_{J^2} - \lambda_{J_z}(\lambda_{J_z} + 1). \tag{15}$$

Since the norm square of any vector in vector space must be non-negative, this requires

$$\lambda_{J^2} \geq |\lambda_{J_z}|(|\lambda_{J_z}| + 1) \geq 0. \tag{16}$$

However, the $J_\pm$ makes λ_{J_z} unbounded unless the vector vanished at some points, i.e. the equality must hold for the upper, and lower bounds of the possible λ_z's. If we replace λ_{J^2} with $j(j + 1)$ where j non-negative. The requirement makes $\lambda_{J_z,\max} = j$, and $\lambda_{J_z,\min} = -j$. Since any pair of λ_{J_z}'s are different by an integer,

$$\lambda_{J_z,\max} - \lambda_{J_z,\min} = 2j \in \mathbb{Z}_0^+. \tag{17}$$

Hence, we can relabel the eigenbasis from $|\lambda_{J^2}, \lambda_{J_z}\rangle$ to $|j, m\rangle$ and choose the proportionality such that

$$
\begin{aligned}
J_+ |j, m\rangle &= \sqrt{j(j+1) - m(m+1)} \, |j, m+1\rangle, \\
J_- |j, m\rangle &= \sqrt{j(j+1) - m(m-1)} \, |j, m-1\rangle,
\end{aligned}
\tag{18}
$$

where $j \in \{0, 1/2, 1, 3/2, ...\}$, and $m \in \{-j, -j+1, ..., j-1, j\}$. With similar method to the proof for J_z, one can directly prove by contradiction that each irreducible representation must have its unique j, i.e., the irreducible representation can be indexed by j. Furthermore, since we have not make any other assumption, j completely identify every irreducible representation of $\mathfrak{su}(2)$, and all irreducible representation with the same j are isomorphic.

5. $SU(2)$'S IRREDUCIBLE REPRESENTATION

Consider a representation of $SU(2)$ that is an exponential of $\mathfrak{su}(2)$'s irreducible representations. Let assume that such representation is reducible. By definition, this means that there is a subspace invariant by any $SU(2)$'s elements. This further implies that it is also invariant by any $\mathfrak{su}(2)$'s. Since corresponding representation of $SU(2)$, and $\mathfrak{su}(2)$ are homomorphic, it entails that there is a subspace of $\mathfrak{su}(2)$'s representation that is invariant to $\mathfrak{su}(2)$, i.e. this specific representation of $\mathfrak{su}(2)$ is reducible, which contradicts the assumption. Therefore, we can directly promote the irreducible representations of $\mathfrak{su}(2)$ to the one for $SU(2)$ by means of matrix exponential along with the same vector space.

6. $SU(2)$'S IMPLEMENTATION

In order to implement the representation the basis of the space in which all group operations would act on need to be carefully selected mainly for the purpose of reducing computational costs. However, it is require the consideration on every symmetry group to decide. Hence, for now, we are assuming the standard $\{|j, m\rangle\}$ basis. In this basis, the irreducible representation with index j is the $2j + 1$ dimensional space representation. Let the representation of a vector is in the ascending order of the basis from $m = -j$ to $m = j$.

Hence,

$$
\begin{aligned}
\left[J_z^{(j)}\right]_{qr} &= (q-j)\delta_{q,r}, \\
\left[J_+^{(j)}\right]_{qr} &= \sqrt{j(j+1)-(q-j)(r-j)}\delta_{q,r+1}, \\
J_-^{(j)} &= J_+^{(j)\top}, \\
J_x^{(j)} &= \left(J_+^{(j)} + J_-^{(j)}\right)/2, \\
J_y^{(j)} &= -i\left(J_+^{(j)} - J_-^{(j)}\right)/2,
\end{aligned}
\tag{19}
$$

with the $u \in SU(2)$ represented as

$$
D_{SU(2)}^{(j)}(u) = \exp\big(i(J_x^{(j)}\phi_x + J_y^{(j)}\phi_y + J_z^{(j)}\phi_z)\big)
\tag{20}
$$

for some $\vec{\phi} \in \mathbb{R}^3$.

7. P'S DEFINITION

P is the group of spatial inversion, i.e. the group consisting of 2 elements: E (identity), and I (inversion) with the only non-trivial multiplication be $II = E$.

8. P'S IRREDUCIBLE REPRESENTATION

From the definition, it is clear that $P \cong C_2$. Hence, there are only two irriducible representations commonly denoted as even, and odd. With similar construction to the $SU(2)$ case, let define the basis (1 dimension) of the two irreducible representations with $|1\rangle$, and $|-1\rangle$, respectively. Therefore,

$$
\begin{aligned}
D_P^{(1)}(E)\,|1\rangle &= |1\rangle\,, \\
D_P^{(1)}(I)\,|1\rangle &= |1\rangle\,, \\
D_P^{(-1)}(E)\,|-1\rangle &= |-1\rangle\,, \\
D_P^{(-1)}(I)\,|-1\rangle &= -\,|-1\rangle\,.
\end{aligned}
\tag{21}
$$

In order to incoperate P to $SU(2)$, the multiplication between these groups need to be considered. At this point, the purely mathematical treatment is not sufficient to know the multiplication behavior, and require physical properties of the systems that the groups are used for. In this case, we are interested in quantum spin systems in which the $SU(2)$ group

is responsible for rotations, and commutes with spatial inversion. This implies that the combined group, i.e. direct product group, $SU(2) \times P$, can be irreducible represented with

$$\left[D_{SU(2)\times P}^{(jp)}(uv)\right]_{qr} = \left[D_{SU(2)}^{(j)}(u)\right]_{qr} D_P^{(p)}(v) \tag{22}$$

for $u \in SU(2)$, $v \in P$, and $p \in \{-1, 1\}$. Furthermore, the vector space basis of this irreducible representation is simply $|j, m, p\rangle = |j, m\rangle |p\rangle$.

9. P'S IMPLEMENTATION

From previous section, with the basis used in $SU(2)$ implementation, one can directly extend it to irreducible representation of $SU(2) \times P$ by adding index p. In this basis, the irreducible representation with index (j, p) is a $2j+1$ dimensional space representation with even inversion if $p = 1$, or odd if $p = -1$. With this representation, for $w = uv \in SU(2) \times P$, $u \in SU(2)$, $v \in P$,

$$D_{SU(2)\times P}^{(jp)}(uv) = D_{SU(2)}^{(j)}(u)D_P^{(p)}(v) \tag{23}$$

with

$$\begin{aligned} D_P^{(p)}(E) &= 1 \\ D_P^{(p)}(I) &= p \end{aligned} \tag{24}$$

10. TR'S DEFINITION

TR is the group of time reversal consisting of E (identity), and θ (time reverse) as generators. However, this group is not isomorphic with P since the physical properties require time-reversal symmetric system to obey Schrödinger's equation. This makes the effect of θ on wavefunction of spinless system be a complex conjugation which means that the group is anti-unitary.

11. TR'S IRREDUCIBLE REPRESENTATION

In corporateing anti-unitary group TR with $SU(2) \times P$ cannot be done in the same fashion as combinding P with $SU(2)$ since the anti-unitary nature makes it impossible to write down common eigenbasis even if θ commutes with every element in $SU(2) \times P$.

To proceed from this point, we need to consider the effects of θ on the basis chosen for $SU(2) \times P$'s implementation. Since the multiplication of two anti-unitary operators is a unitary operator the irreducible representation of $w \in SU(2) \times P$ in $\{\theta \, |j, m, p\rangle\}$ basis is

$$\tilde{D}_{SU(2)\times P}(w) = D_{SU(2)\times P}(\theta^{-1}w\theta)^*. \tag{25}$$

However, similar to P's case, the system in which the $SU(2) \times P$ describes actually makes θ commutes with every element in $SU(2) \times P$. This means that

$$D_{SU(2)\times P}(\theta^{-1}w\theta)^* = D_{SU(2)\times P}(w)^*, \tag{26}$$

but we can always write

$$D_{SU(2)\times P}(w) = \pm\exp\left(i\vec{J} \cdot \vec{\phi}\right) = \pm\exp(i(J_x\phi_x + J_y\phi_y + J_z\phi_z)). \tag{27}$$

This means that

$$D_{SU(2)\times P}(w)^* = \pm\exp(-i(J_x\phi_x - J_y\phi_y + J_z\phi_z)). \tag{28}$$

Then, consider a matrix

$$\left[U^{(j)}\right]_{qr} = i^{2(q-j)}\delta_{q,2j-r}. \tag{29}$$

This matrix is a unitary matrix since

$$\begin{aligned}
\left[U^{(j)\dagger}\right]_{qr} &= (-i)^{2(r-j)}\delta_{r,2j-q}, \\
\left[U^{(j)\dagger}U^{(j)}\right]_{qr} &= \sum_s (-i)^{2(s-j)}\delta_{s,2j-q}i^{2(s-j)}\delta_{s,2j-r} \\
&= \sum_s \delta_{s,2j-q}\delta_{s,2j-r} = \delta_{q,r}, \\
\left[U^{(j)}U^{(j)\dagger}\right]_{qr} &= \sum_s i^{2(q-j)}\delta_{q,2j-s}(-i)^{2(r-j)}\delta_{r,2j-s} \\
&= \sum_s i^{2(q-r)}\delta_{q,2j-s}\delta_{r,2j-s} = i^{2(q-r)}\delta_{q,r} = \delta_{q,r}.
\end{aligned} \tag{30}$$

From this matrix, we can see that

$$\begin{aligned}
\left[U^{(j)\dagger}J_z^{(j)}U^{(j)}\right]_{qr} &= \sum_{s,t}(-i)^{2(s-j)}\delta_{s,2j-q}(s-j)\delta_{s,t}i^{2(t-j)}\delta_{t,2j-r} \\
&= \sum_s (-i)^{2(s-j)}\delta_{s,2j-q}(s-j)\delta_{s,2j-r}i^{2(j-r)} \\
&= (-i)^{2(j-q)}(j-q)\delta_{q,r}i^{2(j-r)} \\
&= -(q-j)\delta(q,r) \\
&= \left[-J_z^{(j)}\right]_{qr},
\end{aligned} \tag{31}$$

$$\left[U^{(j)\dagger}J_+^{(j)}U^{(j)}\right]_{qr} = \sum_{s,t}(-i)^{2(s-j)}\delta_{s,2j-q}\sqrt{j(j+1)-(s-j)(t-j)}\delta_{s,t+1}i^{2(t-j)}\delta_{t,2j-r}$$

$$= \sum_s (-i)^{2(s-j)}\delta_{s,2j-q}\sqrt{j(j+1)-(s-j)(j-r)}\delta_{s,2j-r+1}i^{2(j-r)}$$

$$= (-i)^{2(j-q)}\sqrt{j(j+1)-(j-q)(j-r)}\delta_{2j-q,2j-r+1}i^{2(j-r)} \tag{32}$$

$$= -\sqrt{j(j+1)-(q-j)(r-j)}\delta_{q+1,r}$$

$$= \left[-J_+^{(j)\top}\right]_{qr} = \left[-J_-^{(j)}\right]_{qr},$$

and

$$U^{(j)\dagger}J_-^{(j)}U^{(j)} = U^{(j)\dagger}J_+^{(j)\dagger}U^{(j)}$$

$$= \left(U^{(j)\dagger}J_+^{(j)}U^{(j)}\right)^{\dagger} \tag{33}$$

$$= -J_-^{(j)\dagger} = -J_+^{(j)}.$$

This also directly implies

$$U^{(j)\dagger}J_x^{(j)}U^{(j)} = -J_x^{(j)}$$

$$U^{(j)\dagger}J_y^{(j)}U^{(j)} = J_y^{(j)} \qquad . \tag{34}$$

$$U^{(j)\dagger}J_z^{(j)}U^{(j)} = -J_z^{(j)}$$

Therefore,

$$U^{(j)\dagger}D_{SU(2)\times P}^{(j)}(w)U^{(j)} = \pm U^{(j)\dagger}\exp\left(i\vec{J}^{(j)}\cdot\vec{\phi}\right)U^{(j)}$$

$$= \pm U^{(j)\dagger}\exp\left(i(J_x^{(j)}\phi_x + J_y^{(j)}\phi_y + J_z^{(j)}\phi_z)\right)U^{(j)}$$

$$= \pm\exp\left(-i(J_x^{(j)}\phi_x - J_y^{(j)}\phi_y + J_z^{(j)}\phi_z)\right) \tag{35}$$

$$= D_{SU(2)\times P}^{(j)}(w)^*$$

$$= \tilde{D}_{SU(2)\times P}^{(j)}(w).$$

This directly means that the $SU(2) \times P$'s irreducible representation in $\{|j,m,p\rangle\}$, and $\{\theta\,|j,m,p\rangle\}$ basis are equivalent. This implies that the corepresentation of $SU(2) \times P \times TR$ for $w \in SU(2) \times P$ is

$$D_{SU(2)\times P\times TR}^{(j)}(w) = \begin{bmatrix} D_{SU(2)\times P}^{(j)}(w) & 0 \\ 0 & U^{(j)\dagger}D_{SU(2)\times P}^{(j)}(w)U^{(j)} \end{bmatrix} \tag{36}$$

where the first half of the basis is $\{|j,m,p\rangle\}$ while the second half is $\{\theta\,|j,m,p\rangle\}$.

In order to get the corepresentation for θ, we know that, since $\theta^{-1}w\theta$, θ^{-2} and θ^2 are unitary,

$$\tilde{D}_{SU(2)\times P}^{(j)}(\theta^{-1}w\theta) = D_{SU(2)\times P}^{(j)}(\theta^{-2}w\theta^2)^*. \tag{37}$$

However, this exact thing is also equal

$$\tilde{D}^{(j)}_{SU(2)\times P}(\theta^{-1}w\theta) = U^{(j)\dagger}D^{(j)}_{SU(2)\times P}(\theta^{-1}w\theta)U^{(j)}. \tag{38}$$

Hence,

$$D^{(j)}_{SU(2)\times P}(\theta^{-2}w\theta^2) = U^{(j)\dagger*}D^{(j)}_{SU(2)\times P}(\theta^{-1}w\theta)^*U^{(j)*}$$

$$D^{(j)}_{SU(2)\times P}(\theta^2)^{-1}D^{(j)}_{SU(2)\times P}(w)D^{(j)}_{SU(2)\times P}(\theta^2) = U^{(j)\dagger*}U^{(j)\dagger}D^{(j)}_{SU(2)\times P}(w)U^{(j)}U^{(j)*} \tag{39}$$

$$U^{(j)}U^{(j)*}D^{(j)}_{SU(2)\times P}(\theta^2)^{-1}D^{(j)}_{SU(2)\times P}(w) = D^{(j)}_{SU(2)\times P}(w)U^{(j)}U^{(j)*}D^{(j)}_{SU(2)\times P}(\theta^2)^{-1}.$$

From Schur's lemma, since $D^{(j)}_{SU(2)\times P}$ is an irreducible representation, there is a constant $\lambda \in \mathbb{C}$ such that

$$tU^{(j)}U^{(j)*} = D^{(j)}_{SU(2)\times P}(\theta^2). \tag{40}$$

With similar process but substitute $w = \theta^2$,

$$D^{(j)}_{SU(2)\times P}(\theta^2)^* = U^{(j)\dagger}D^{(j)}_{SU(2)\times P}(\theta^2)U^{(j)}. \tag{41}$$

This, together with the previous one, gives

$$t^*U^{(j)}U^{(j)*} = D^{(j)}_{SU(2)\times P}(\theta^2). \tag{42}$$

Hence, $t \in \mathbb{R}$. Moreover, the unitarity of $U^{(j)}$, and $SU(2) \times P$'s irreducible representation, restricts $t = \pm 1$. With all these results, one now can write the corepresentation of $SU(2) \times P \times TR$ for $\theta \in T$ as

$$D^{(j)}_{SU(2)\times P\times TR}(\theta) = \begin{bmatrix} 0 & D^{(j)}_{SU(2)\times P}(\theta^2) \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & tU^{(j)}U^{(j)*} \\ 1 & 0 \end{bmatrix}. \tag{43}$$

One can further simplify the representation by changing the basis $\{\theta\,|j,m,p\rangle\}$ with $U^{(j)}$,

$$D^{(j)}_{SU(2)\times P\times TR}(w) = \begin{bmatrix} D^{(j)}_{SU(2)\times P}(w) & 0 \\ 0 & D^{(j)}_{SU(2)\times P}(w) \end{bmatrix}$$

$$D^{(j)}_{SU(2)\times P\times TR}(\theta) = \begin{bmatrix} 0 & tU^{(j)} \\ U^{(j)} & 0 \end{bmatrix}. \tag{44}$$

12. TR'S IMPLEMENTATION