Representation theory for Geometric Quantum Machine Learning

Michael Ragone
joint work at Los Alamos National Laboratory with Paulo Braccia, Quynh T. Nguyen, Louis Schatzki, Patrick J. Coles, Frédéric Sauvage, Martín Larocca, and M. Cerezo

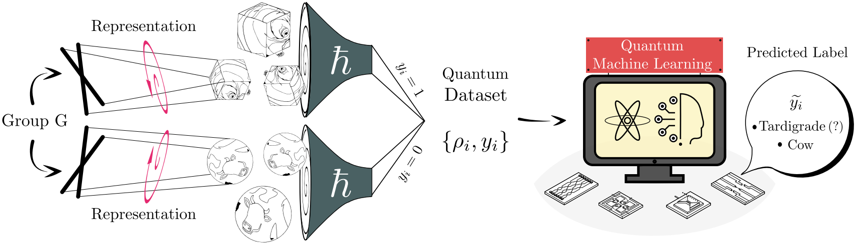
arXiv:2210.07980

University of California, Davis

January 21, 2023

Why symmetry?

- *import: Marco Cerezo's QML Talk*
- Key QML problem: barren plateaus!
- Symmetry yields powerful inductive biases, tightly constraining search space size and structure: "a rotated tardigrade is still a tardigrade, and a rotated cow is still a cow".



Why symmetry?

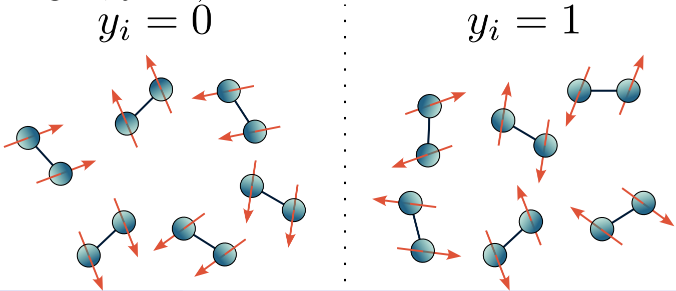
- Morally, symmetry describes "transformations which leave some property invariant".
- For QML, this means *label invariance* under some transformations of the data.
- Mathematically, symmetry is implemented by group representations.
  - Groups capture symmetry at an abstract level.
  - Representations describe how a group "acts" on a vector space, often via unitaries on a Hilbert space. This can be thought of as a more physical symmetry.
- We will define these after setting up our main example: quantum phase classification.

Ex: quantum phase classification

Heisenberg XXX chain. Local Hilbert spaces are qubits $\mathbb{C}^2$, and $J \in \mathbb{R}$ a parameter:

$$H = J \sum_{j=1}^{n-1} X_j X_{j+1} + Y_j Y_{j+1} + Z_j Z_{j+1},$$

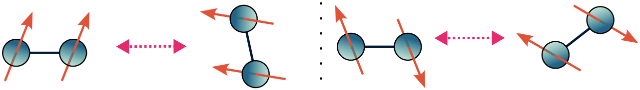
where X, Y, Z are Pauli matrices. Varying J, we get ground states which are either highly aligned (ferromagnet, $y_i = 0$) or anti-aligned (antiferromagnet, $y_i = 1$).



Ex: quantum phase classification

We can use a QML framework, like a quantum neural network, to label ground states as ferromagnets or antiferromagnets. Let $f : \mathcal{R} \to \{0, 1\}$ be a function labeling quantum states.

- *Data:* pure states $|\psi_i\rangle \in \mathcal{R}$
- *Labels:* Ferromagnetic $y_i = 0$ and antiferromagnetic $y_i = 1$
- Let's think about the structure of this problem.
  - Swapping two particles preserves alignment.
  - Identically rotating each particle preserves alignment.
- These operations may change the data, but the labels are invariant! This is a *symmetry*.



Ex: discrete symmetry

"Swapping two particles doesn't change ferromagnetism."

- Group: $G = \mathbb{Z}_2 = \{1, -1\}$ with multiplication.
- Representation space: $V = \mathbb{C}^2 \otimes \mathbb{C}^2$ (2 qubits)
- Representation R: each group element $g \in G$ gets a corresponding linear map $R_g : V \to V$,

$$R_1 \cdot |\psi_i\rangle = |\psi_i\rangle \,,$$
$$R_{-1} \cdot |\psi_i\rangle = \texttt{SWAP}\, |\psi_i\rangle \,.$$

Symmetry = Label invariance under group representation: for any quantum state $\rho_i \in \mathcal{R} \subseteq V$,

$$f(|\psi_i\rangle) = f(R_g \cdot |\psi_i\rangle), \quad \text{for all } g \in G.$$

So, $f(|\psi_i\rangle) = f(|\psi_i\rangle)$ and $f(|\psi_i\rangle) = f(\texttt{SWAP}\, |\psi_i\rangle)$.

Ex: continuous symmetry

"Rotated ferromagnets are still ferromagnets."

- Group: $G = SU(2)$
- Representation space: $V = \mathbb{C}^2 \otimes \mathbb{C}^2$ (2 qubits)
- Representation: $R_g = g^{\otimes 2}$
  - e.g. if $g = e^{itX} \in SU(2)$, then its representative $R_g : V \to V$ is

$$R_g \cdot |\psi_i\rangle = (e^{itX} \otimes e^{itX}) |\psi_i\rangle .$$

Symmetry = Label invariance under rep R of $G = SU(2)$:

$$\begin{aligned} f(|\psi_i\rangle) &= f(R_g \cdot |\psi_i\rangle) \\ &= f(g^{\otimes 2} |\psi_i\rangle). \end{aligned}$$

This is a huge inductive bias! We will eventually see that an 3D manifold of states is now identified as a single unit.

The formalities: groups

- Keep the example of rotations $G = SU(2)$ in mind, where composition (matrix multiplication) is the group operation.
- A *group* G is a set with a binary operation obeying some axioms.
  - *Binary operation:* If $g, h \in G$, then $gh \in G$. (Rotation by $h = e^{iY}$ and then rotation by $g = e^{iX}$ is the same as a rotation by $gh = e^{iX}e^{iY}$.)
  - *Identity:* There's a $\mathbb{1} \in G$, such that $g = g\mathbb{1} = \mathbb{1}g$ for any $g \in G$. (The identity matrix $e^0 = \mathbb{1} \in SU(2)$ is a rotation by 0 degrees)
  - *Inverse:* For every $g \in G$, there exists a g^{-1} with $\mathbb{1} = gg^{-1} = g^{-1}g$. (Every rotation $g = e^{i\theta X}$ has a reverse rotation $g^{-1} = e^{-i\theta X}$)
  - *Associativity:* if $g, h, k \in G$, then

    $$g(hk) = (gh)k.$$

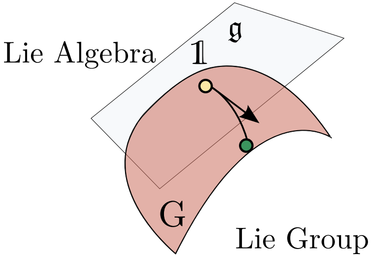
    (You can be lazy with order of operations)

Discrete and continuous groups

$G = SU(2)$ is a *Lie group* or *continuous group*, meaning it's also a manifold (we can parameterize it with coordinates).[1]

- $\implies$ We can draw continuous paths in G and take derivatives along these paths.
- Ex: time evolution with infinitesimal generator $iX \in \mathfrak{su}(2)$:

$$\{U_t := e^{itX} \mid t \in \mathbb{R}\} \text{ is a path in } SU(2)$$



[1] Matrix groups, like $SO(n)$ or $GL(n)$, are often Lie groups, and many commonly appearing Lie groups are matrix groups.

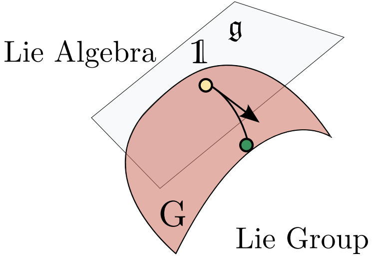
Discrete and continuous groups

Taking directional derivatives along paths in G gives a vector space with matrix commutator $[\cdot,\cdot]$: this is a *Lie algebra* $\mathfrak{g}$.

- Ex: Take the path of rotations $\{U_t = e^{itX}\}$. The infinitesimal generator is then

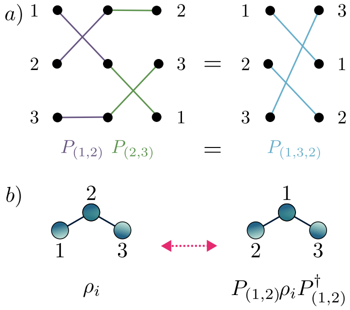
$$\frac{d}{dt}U_t\Big|_{t=0} = iXe^0 = iX \in \mathfrak{su}(2).$$

  In fact, Paulis X, Y, Z form a basis for $\mathfrak{su}(2)$, and the commutation relations tell us about its Lie algebra structure!



Discrete and continuous groups

- If the underlying set is discrete, we have a *discrete group*.
- Key example: the symmetric group S_n of permutations on n letters.
  - Two-particle swap is a special case: $S_2 \cong \mathbb{Z}_2 = \{1, -1\}$.
  - The symmetric group commonly arises in quantum as *tensor index permutations*, like qubit swaps.



Representation theory

- A *representation* R of a group G on a vector space V is a homomorphism $R : G \to \mathrm{GL}(V)$, the general linear group (invertible matrices on V)
  - *Homomorphism* just means that this map respects the group structure: if $g, h \in G$ and $V = \mathbb{C}^n$, then R_g, R_h are $n \times n$ complex matrices obeying
$$R_g R_h = R_{gh}.$$
- The first and most boring example: the trivial representation, where $V = \mathbb{C}$ and
$$R_g = \mathbb{1}, \quad \text{for all } g \in G.$$
Sanity check: if $G = \mathbb{Z}_2$, and $g = 1, h = -1$, then
$$\mathbb{1} = R_{-1} = R_1 R_{-1} = \mathbb{1}\mathbb{1} = \mathbb{1}.$$

Representation theory: $G = \mathbb{Z}_2$

- Let $V = \mathbb{C}$ and define the representation

$$R_1 = \mathbb{1}, R_{-1} = -\mathbb{1}.$$

Sanity check:

$$\mathbb{1} = R_1 = R_{-1}R_{-1} = (-\mathbb{1})(-\mathbb{1}) = \mathbb{1}.$$

Seems to work! Let's revisit the earlier two-qubit swap example.

Representation theory: $G = \mathbb{Z}_2$

- Let $V = \mathbb{C}^2 \otimes \mathbb{C}^2$ and recall:

$$R_1 = \mathbb{1}$$
$$R_{-1} = \texttt{SWAP}.$$

Sanity check: if $|\psi_i\rangle \in V$, then

$$\begin{aligned} |\psi_i\rangle = R_1 \cdot |\psi_i\rangle &= (R_{-1} \cdot R_{-1}) \cdot |\psi_i\rangle \\ &= R_{-1} \cdot (\texttt{SWAP}\,|\psi_i\rangle) \\ &= \texttt{SWAP}(\texttt{SWAP}\,|\psi_i\rangle) \\ &= |\psi_i\rangle \end{aligned}$$

- Onto the continuous examples! Same group laws, but more structure because we can take derivatives to pass to the Lie algebra $\mathfrak{g}$.

Representation theory: $G = SU(2)$

- *Defining representation:*[2] This is the rep where the representative matrices are the same as in the group itself.
  - If $G = SU(2)$, then let $V = \mathbb{C}^2$ and define the rep $U : V \to V$ to be $U_g := g$. We often write U for unitary reps.
- We can build new representations by *tensoring* $\otimes$ and *direct summing* $\oplus$. Let's revisit the ferromagnet example and do some rep theory to see what it can do for us!
- Two identically rotating qubits: $V = \mathbb{C}^2 \otimes \mathbb{C}^2$, and the rep is $g \mapsto U_g \otimes U_g$.
- Now, one of the most useful lemmas in the business.

[2] Physicists often call this the fundamental representation. Mathematicians have a slightly wider definition for "fundamental reps" which includes the defining rep.

A powerful lemma

Simultaneous Block-Diagonalization:
Let $U : G \to GL(V)$ a rep of G and $H = H^*$ such that

$$[U_g, H] = 0 \quad \text{for all } g \in G.$$

Then, for any eigenvector $|\psi\rangle$ with eigenvalue λ, $U_g |\psi\rangle$ is also an eigenvector of H with eigenvalue λ.
Proof:

$$H(U_g |\psi\rangle) = U_g H |\psi\rangle = U_g \lambda |\psi\rangle = \lambda(U_g |\psi\rangle).$$

Tensor rep of $G = SU(2)$

- It's pretty easy to see that for all $g \in SU(2)$,

$$[U_g \otimes U_g, \texttt{SWAP}] = 0.$$

$\implies$ Use the lemma! Then, diagonalizing $\texttt{SWAP}$, which acts as $\mathbb{1}$ on the symmetric subspace $\texttt{Sym}$ spanned by $\{|11\rangle, |10\rangle + |01\rangle, |00\rangle\}$ and $-\mathbb{1}$ on the antisymmetric subspace $\texttt{Alt}$ spanned by $\{|10\rangle - |01\rangle\}$, gives a block-diagonal structure for the representative matrices $U_g \otimes U_g$:

$$U_g \otimes U_g = \begin{pmatrix} \multicolumn{3}{c}{\boxed{}} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & \Box \end{pmatrix}, \quad g \in SU(2).$$

- These are *invariant subspaces* of the rep space $V = \texttt{Sym} \oplus \texttt{Alt}$. Are they the smallest invariant subspaces, aka *irreducible*?

Reducibility and the Lie algebra

- `Alt` is 1 dimensional, so irreducible.
- Theme: Lie groups are tricky, but their Lie algebras (directional derivatives) are vector spaces and easier to work with.
- *Theorem:* a rep space V is irreducible under a Lie group G if and only if it is irreducible under its Lie algebra $\mathfrak{g}$.
- So, we can work with the directional derivatives, i.e. the Paulis $\mathfrak{su}(2)$. For instance, $e^{itZ} \in SU(2)$, so passing to a representation of $\mathfrak{su}(2)$ on V by taking derivatives,

$$-i\frac{d}{dt}(e^{itZ} \otimes e^{itZ})\Big|_{t=0} = Z \otimes \mathbb{1} + \mathbb{1} \otimes Z.$$

  With these, it will be much easier to see that there's no invariant subspace under the Lie algebra representation.

Angular momentum operators = Lie algebra representation of $\mathfrak{su}(2)$

Work on the basis $\{|11\rangle, |10\rangle + |01\rangle, |00\rangle\}$ of `Sym`:

$$(Z \otimes \mathbb{1} + \mathbb{1} \otimes Z)|11\rangle = -2|11\rangle$$
$$(Z \otimes \mathbb{1} + \mathbb{1} \otimes Z)(|10\rangle + |01\rangle) = 0(|10\rangle + |01\rangle)$$
$$(Z \otimes \mathbb{1} + \mathbb{1} \otimes Z)|00\rangle = 2|00\rangle$$

Define "creation" and "annihilation" operators $\widetilde{\sigma}^{\pm} = \sigma^{\pm} \otimes \mathbb{1} + \mathbb{1} \otimes \sigma^{\pm}$ in the image of this $\mathfrak{su}(2)$ representation where $\sigma^{\pm} = \frac{1}{2}(X \mp iY)$, i.e.

$$|0\rangle \xmapsto{\sigma^{+}\cdot} |1\rangle \xmapsto{\sigma^{+}\cdot} 0, \qquad |1\rangle \xmapsto{\sigma^{-}\cdot} |0\rangle \xmapsto{\sigma^{-}\cdot} 0.$$

No invariant subspace, so `Sym` an irreducible representation!

$$|11\rangle \xmapsto{\tilde{\sigma}^{-}\cdot} \big(|10\rangle + |01\rangle\big) \xmapsto{\tilde{\sigma}^{-}\cdot} 2|00\rangle$$
$$|00\rangle \xmapsto{\tilde{\sigma}^{+}\cdot} \big(|10\rangle + |01\rangle\big) \xmapsto{\tilde{\sigma}^{+}\cdot} 2|11\rangle .$$

The irreducibility payoff

- Back to the ferromagnet: can calculate (using Lie algebra) that the Hamiltonian H has

$$[U_g \otimes U_g, H] = 0.$$

$\implies$ Simultaneous Block-Diagonalization! $V = \texttt{Sym} \oplus \texttt{Alt}$.

- You cleverly calculate for some J that $|00\rangle$ is a ground state of energy λ. But $|00\rangle \in \texttt{Sym}$, and $\texttt{Sym}$ is irreducible, so you immediately know *all* of $\texttt{Sym}$ are ground states!
- "Rotated ferromagnet is still a ferromagnet" $\implies$ Our state labels are irreducible representations of $SU(2)$!

$$\texttt{Sym} \iff \text{ferromagnet}, \qquad \texttt{Alt} \iff \text{antiferromagnet}.$$

What just happened?

- We saw a symmetry in the problem: "rotated ferromagnets are still ferromagnets".
  - In other words, the Heisenberg Hamiltonian H is invariant under this symmetry and by the lemma, the symmetry respects the energy labels.
- The symmetry is realized by the tensor representation $g \mapsto U_g \otimes U_g$ of $G = SU(2)$ on two qubits $V = \mathbb{C}^2 \otimes \mathbb{C}^2$.
- The representation $U_g \otimes U_g$ can be broken into a block-diagonal structure $V = \texttt{Sym} \oplus \texttt{Alt}$, and there is no smaller block-diagonal structure (irreducible).
  - Key was passing from Lie Group G to Lie algebra $\mathfrak{g}$, since linear things are easier.
- Irreducibility cut down number of possible states to explore from 4 dimensions to 2, and "ferro/antiferro" labels happened to coincide with irreducibles.

Just the beginning

- The symmetry program has been one of the most successful programs in science, across quantum field theory, condensed matter, classical machine learning...A key reason is that *symmetry gives block diagonal structures which reduce problems.*
- In QML, the natural extension is to look at symmetry transformations on data spaces which leave labels invariant. This gives inductive biases, which *reduces problems.*
- Exciting evidence: in arXiv:2210.09974[3], it was analytically shown that permutation-equivariant QNNs "do not suffer from barren plateaus, quickly reach overparametrization, and generalize well from small amounts of data".

[3] Schatzki, Larocca, Nguyen, Sauvage, Cerezo 2022 preprint

Just the beginning

- There are many QML tasks (and other quantum algorithms) which possess natural symmetry that is not being exploited: these may make difficult analysis and numerics problems tractable!
- When data has a symmetry (and the algorithm commutes with this symmetry[4]), we have an inductive bias, which may be able to reduce circuit and sample complexities, as well as kill barren plateaus and other trainability hurdles.
- My favorite example: unsupervised learning of ground state quantum phase diagrams for models with more interesting symmetries, like other Lie types $SO(n), Sp(n) \ldots$.

[4] When a function commutes with a symmetry, this is called *equivariance*.