On Representations of Compact Groups over Fields of Characteristic Zero

Owen L. Howell[1]

[1]*Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Ave., Boston, MA 02115, USA*

(Dated: January 3, 2025)

This is a set of notes dealing with the representations of compact groups over the field $\mathbb{R}$ or $\mathbb{C}$. The key ideas of representation theory are covered; namely Schur's lemma, the orthogonality theorems, the Peter-Weyl theorem and the Parseval-Plancherel theorem. We discuss some operations that can be performed on representations, including the direct sum, induced/restriction and tensor product representations. Armed with these tools, we discuss a few specific examples of group representations that occur frequently in physics and deep learning. We also discuss a few more 'exotic' topics that are not usually covered in representation theory textbooks. We focus on the representation theory of the symmetric group S_n and the Schur-Weyl duality. Finally, we review Bochner's theorems on commutative and non-commutative groups with applications to kernel methods.

Contents

I. INTRODUCTION

Representation theory is the mathematical framework for studying abstract groups by representing their elements as linear transformations (or matrices) on vector spaces. The representation theory of compact groups is particularly elegant, and is a powerful generalization of standard Fourier analysis. The main tool in representation theory of compact groups is the decomposition of the group actions into simpler, irreducible representations; analogous to breaking down complex waveforms into sinusoidal components in Fourier analysis. Methods in representation theory can be utilized anytime a system exhibits symmetry or can be described by a group action. These methods allow us to analyze and simplify problems across mathematics, physics, and engineering by exploiting the structure of underlying symmetry.

A. Who Cares? Why Should I Read This?

As mentioned previously, there exist many excellent textbooks on group theory [6, 13, 27, 31]. These books were published before the advent of equivarient deep learning [5]. I hope that these notes provide an overview of representation theory on compact groups which includes both some of the recent developments in equivariant learning theory and the historical uses of representation theory in physics. One very fruitful research direction has been to take classical representation theory and apply it to machine learning problems [9, 15, 24]. The goal of these notes is to give the reader a set of tools that can potentially be applied to equivarient deep learning research, while also discussing some of the historical development of the subject.

B. History of Representation Theory

The representation theory of compact groups has its roots in 19th-century harmonic analysis and the study of symmetry in physics and mathematics. The history of the representation theory of compact groups can be split into two eras: Commutative and Non-Commutative. The Commutative era began with the development of classical electromagnetism and Joseph Fourier's discovery of Fourier Analysis in 1807. The idea of decomposing functions into periodic components led to the representation theory of the circle group (also called $U(1)$ or $SO(2)$) and its finite subgroups (namely $\mathbb{Z}_N$ the cyclic groups of order N). The non-commutative era was mostly motivated by the development of quantum mechanics in the 1920s. In quantum mechanics, symmetries act on Hilbert space states to form representations. For this reason, projective representations over the projective complex space. The complex representations of groups like $SU(2)$ or $SU(5)$ form the backbone of many results in quantum chemistry and particle physics. In the early 20th century, Hermann Weyl established a foundation for compact groups, proving that every finite-dimensional representation of a compact group can be decomposed into irreducible representations.

C. Applications of Representation Theory

We summarize some applications of representation theory of compact groups.

Physics and Chemistry Compact groups, such as $SU(2)$ and $SU(3)$, play a critical role in quantum mechanics and particle physics, describing spin, angular momentum, and the Standard Model's gauge symmetries. Molecular symmetry groups are compact, and their representations explain spectral lines, bonding, and reaction mechanisms.

Harmonic Analysis, Topology and Geometry In algebraic topology, symmetry groups are used to study invariants of spaces. Compact Lie groups, feature prominently in the study of fiber bundles and gauge theory, connecting topology with quantum field theory. Representations of these groups help classify principal bundles and elucidate the structure of topological spaces through their symmetries.

Engineering and Signal Processing Representation theory has a broad impact in engineering and computer graphics. Spherical harmonics efficiently encode 3D models by capturing their rotational symmetries. Spherical harmonics are used in surface reconstruction, lighting models, and shape recognition in 3D imaging.

Equivariant Machine Learning Recent applications include equivariant neural networks, leveraging compact groups to build models invariant to symmetries of the problem instance.

In summary, the representation theory of compact groups bridges abstract algebra, geometry, and analysis with far-reaching implications across both pure and applied sciences. It continues to be a dynamic area of research with modern applications in deep learning and theoretical physics. These note provide a good introduction to some of the techniques and ideas that I have found useful. These notes are a compilation of my own work [4, 14–16, 33]. I emphasize that these notes are not a comprehensive exposition of representation theory and if you are interested in learning more, please see [6, 13, 27, 31].

II. GROUP THEORY

We establish some notation and review some elements of representation theory. For a comprehensive review of representation theory, please see [26, 31]. The identity element of any group G will be denoted as e. A subgroup H of G will be denoted as $H \subseteq G$. We will always work over the field $\mathbb{C}$ unless otherwise specified. A group G is a non-empty set combined with a associative binary operation $\cdot : G \times G \to G$ that satisfies the following properties

$$\begin{aligned}&\text{existence of identity: } e \in G, \text{ s.t. } \forall g \in G, \;\; e \cdot g = g \cdot e = g \\ &\text{existence of inverse: } \forall g \in G, \implies \exists g^{-1} \in G, \;\; g \cdot g^{-1} = g^{-1} \cdot g = e\end{aligned}$$

Oftentimes, we wish to work with a group that can acts on a set of objects in a natural way. This is formalized with the concept of a *group action*.

A. Group Actions

Let Ω be a set. A group action Φ of G on Ω is a map $\Phi : G \times \Omega \to \Omega$ which satisfies

$$\begin{aligned}&\text{Identity Property: } \forall \omega \in \Omega, \quad \Phi(e, \omega) = \omega \\ &\text{Compositional Property: } \forall g_1, g_2 \in G, \;\; \forall \omega \in \Omega, \quad \Phi(g_1 g_2, \omega) = \Phi(g_1, \Phi(g_2, \omega))\end{aligned} \tag{1}$$

We will often suppress the Φ function and write $\Phi(g, \omega) = g \cdot \omega$.

$$\begin{tikzcd} \Omega \arrow[r, "\Psi"] \arrow[d, "\Phi(g,\cdot)"] & \Omega' \arrow[d, "\Phi'(g,\cdot)"] \\ \Omega \arrow[r, "\Psi"] & \Omega' \end{tikzcd}$$

FIG. 1: Commutative Diagram For G-equivariant function: Let $\Phi(g, \cdot) : G \times \Omega \to \Omega$ denote the action of G on Ω. Let $\Phi'(g, \cdot) : G \times \Omega' \to \Omega'$ denote the action of G on Ω'. The map $\Psi : \Omega \to \Omega'$ is G-equivariant if and only if the following diagram is commutative for all $g \in G$.

Let G have group action Φ on Ω and group action Φ' on Ω'. A mapping $\Psi : \Omega \to \Omega'$ is said to be G-equivariant if and only if

$$\forall g \in G, \forall \omega \in \Omega, \quad \Psi(\Phi(g, \omega)) = \Phi'(g, \Psi(\omega)) \tag{2}$$

Diagrammatically, the map Ψ is G-equivariant if and only if the diagram II A is commutative.

B. Lie Groups

Lie group theory is the study of continuous groups. We review some basic concepts of Lie group theory. A full treatment of Lie group theory can be found in [11, 13, 27, 31]. A Lie group G is a group that is also a smooth manifold with the requirement that, for all $g, h \in G$, the map $g \times h \to gh : G \times G \to G$ is smooth and the map $g \to g^{-1} : G \to G$ is smooth. A homeomorphism of Lie groups is a smooth map $\Phi : G \to H$ that satisfies the relation

$$\forall g, g' \in G, \quad \Phi(gg') = \Phi(g)\Phi(g')$$

The Haar measure [28], is the volume element dg of the Lie group G which is left invariant, we have that

$$\forall h \in H, \quad \int_{g \in G} d(hg) = \int_{g \in G} dg$$

For compact groups, the Haar measure is both left and right invariant so that $d(hg) = dg = d(gh)$. Left and right invariance uniquely defines the Haar measure dg on a compact group. Homogeneous spaces $X = G/H$ of G inherit a measure dx on X from the Haar measure. The space of volume elements on X is one-dimensional so

$$\forall g \in G, \quad d(g \cdot x) = \Delta(g^{-1})dx$$

must hold where $\Delta(g) : G \to \mathbb{C}$ is called the modular function of the volume element dx. A famous result of (cite) states that all compact groups are unimodular. In the unimodular case the volume element dx is called the invariant measure on G.

1. *Lie Algebra*

A Lie algebra $\mathfrak{g}$ is a vector space equipped with a anti-symmetric two-form $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which satisfies the Jacobi identity,

$$\text{Jacobi: } [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

A homeomorphism of Lie algebras is a map $\phi : \mathfrak{g} \to \mathfrak{h}$ that preserves the Lie bracket of $\mathfrak{g}$ so that

$$\forall X, Y \in \mathfrak{g}, \quad \phi([X,Y]) = [\phi(X), \phi(Y)]$$

Let X_i be a basis of the Lie algebra $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is called semi-simple if there is no proper subset J_i of the X_i such that the J_i are an idea of $\mathfrak{g}$ under the Lie bracket operator $[\cdot,\cdot]$. Let X_i be a basis of the Lie algebra $\mathfrak{g}$. The structure constants f_{ij}^k of $\mathfrak{g}$ are defined as

$$[X_i, X_j] = \sum_j f_{ij}^k X_k$$

so that the constants f_{ij}^k are the decomposition of the Lie bracket in the vector space $\mathfrak{g}$.

III. REPRESENTATION THEORY

Let V be a vector space over the field $\mathbb{C}$. A complex representation (ρ, V) of a group G consists of the vector space V and a group homomorphism $\rho : G \to \text{Hom}[V,V]$. By definition, the homomorphism ρ must satisfy

$$\forall g, g' \in G, \ \ \forall v \in V, \ \ \rho(g)\rho(g')v = \rho(gg')v$$

Heuristically, a group representation can be thought of as the embedding of an group (which is an abstract mathematical object) into a set of matrices (which we as computer scientists like because matrices are naturally stored as arrays!). Two representations (ρ, V) and (σ, W) are said to be equivalent representations if there exists a matrix Φ

$$\forall g \in G, \quad \Phi\rho(g) = \sigma(g)\Phi$$

The linear map Φ is said to be a G-intertwiner of the (ρ, V) and (σ, W) representations. The space of all G-intertwiners is denoted as $\text{Hom}_G[(\rho, V), (\sigma, W)]$. Specifically,

$$\text{Hom}_G[(\rho, V), (\sigma, W)] = \{ \ \Phi : V \to W \ \mid \ \forall g \in G, \ \ \Phi\rho(g) = \sigma(g)\Phi, \ \ \Phi \text{ is linear } \}$$

The sum of two G-intertwiners is again G-intertwiner and $\text{Hom}_G[(\rho, V), (\sigma, W)]$ forms a vector space over $\mathbb{C}$. The vector space of of G-intertwiners from a representation to itself is called the G endomorpism space of the representation (ρ, V),

$$\text{End}_G[(\rho, V)] = \text{Hom}_G[(\rho, V), (\rho, V)]$$

which we will refer to as the *endomorpism space* of (ρ, V). Much of classical group theory studies the structure of the intertwiners of representations [6]. A representation (ρ, V) is said to be a unitary representation if the vector space V can be equipped with an inner product $\langle \cdot, \cdot \rangle$ such that

$$\forall g \in G, \ \ \forall v, w \in V, \quad \langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle$$

The unitary theorem in representation theory [6] says that any representation of a compact group G is equivalent to a unitary representation of G. A representation is said to be reducible if it breaks into a direct sum of smaller representations. Specifically, a unitary representation ρ is reducible if there exists an unitary matrix U such that

$$\forall g \in G, \quad \rho(g) = U[\bigoplus_{i=1}^{k} \sigma_i(g)]U^\dagger$$

where $k \geq 2$ and σ_i are smaller representations of G. The set of all non-equivalent unitary representations of a group G will be denoted as $\hat{G}$. All representations of compact groups G can be decomposed into direct sums of irreducible representations. Specifically, if (σ, V) is a G-representation,

$$(\sigma, V) = U[\bigoplus_{\rho \in \hat{G}} m_\sigma^\rho (\rho, V_\rho)]U^\dagger$$

where U is a unitary matrix and the integers m_σ^ρ denote the number of copies of the irreducible (ρ, V_ρ) in the representation (σ, V).

A. Lie Group and Lie Algebra Representations

Representations of Lie groups are defined in the same way as representations of finite groups. Let V be a vector space. A representation of a Lie group is a Lie group homeomorpism $\rho : G \to GL(V)$ and a vector space V satisfying,

$$\forall g \in G, \ \forall v \in V, \quad \rho(gg')v = \rho(g)\rho(g')v$$

We can similarly speak of a Lie algebra representation as a homeopmorpism $\sigma : \mathfrak{g} \to GL(V)$ that preserves Lie bracket structure

$$\forall X, Y \in \mathfrak{g}, \quad \sigma([X, Y]) = [\sigma(X), \sigma(Y)]$$

If G is a connected group, the map $\exp : \mathfrak{g} \to G$, is defined as

$$\forall X \in \mathfrak{g}, \quad \exp(itX) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} X^n$$

The key property of exp is that the exponential map exp commutes with homeomorphism of algebra and group III A, so that there is an isomerism between Lie algebra representations and Lie group representations.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "d\Phi|_e"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"'] \\ G \arrow[r, "\Phi"] & H \end{tikzcd}$$

FIG. 2: The exponential map: Let $\Phi : G \to H$ be a homeomorpism of groups. Let $d\Phi|_e : \mathfrak{g} \to \mathfrak{h}$ be the derivative map evaluated at the identity of G. Then, the above map is commutative.

1. Adjoint Representation

a. Little Adjoint ad Representation The adjoint (sometime called the little adjoint) ad representation is a canonical representation of a Lie algebra. The adjoint action is defined via the formula

$$\mathrm{ad}(X)Y = [X, Y]$$

So that as a matrix the ad representation is of dimension equal to the number of Lie algebra basis elements. The adjoint action satisfies

$$[\mathrm{ad}(X), \mathrm{ad}(Y)] = \mathrm{ad}([X, Y])$$

which preserves the Lie bracket structure and is thus a valid Lie algebra representation. The adjoint action acts directly on $\hat{g}$, and the dimension of the adjoint representation is the dimension of the vector space $\hat{g}$.

b. Big Adjoint Ad Representation There is an analogous adjoint (sometimes called the big adjoint) Ad representation of the Lie group G on $\mathfrak{g}$. Consider the conjugation map $\Phi_g : G \to G$ on the Lie group G given by

$$\Phi_g(h) = ghg^{-1}$$

the conjugation map is an Lie automorpism of G. The adjoint map Ad_g evaluated at $g \in G$ is then the conjugation map evaluated at the identity

$$\forall g \in G, \quad Ad_g = d\Phi_g|_e : T_e(G) \to T_e(G)$$

so that for fixed $g \in G$, $Ad_g : \mathfrak{g} \to \mathfrak{g}$. Thus, $Ad_g : G \to \mathrm{aut}(\mathfrak{g})$. Let $X \in \mathfrak{g}$,

$$\forall g \in G, \quad Ad_g X = \frac{d}{dt}[g \exp(tX) g^{-1}]|_{t=0}$$

Note that

$$\forall g, g' \in G, \quad Ad_g \circ Ad_{g'} = Ad_{gg'}$$

so that $(Ad, \mathfrak{g})$ is a Lie group representation of G with dimension equal to the vector space dimension of $\mathfrak{g}$. Let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathfrak{g}$. The inner product $\langle \cdot, \cdot \rangle$ is said to be Ad-invariant if and only if,

$$\forall g \in G, \ \forall x, y \in \mathfrak{g}, \quad \langle x, y \rangle = \langle Ad_g x, Ad_g y \rangle$$

IV. SCHUR'S LEMMA

Schur's lemma is one of the fundamental results in representation theory [31]. Let G be a compact group. Let (ρ, V) and (σ, W) be irreducible representations of G. Then, Schur's lemma states the following: Let $\Phi : V \to W$ be an intertwiner of (ρ, V) and (σ, W). Then, Φ is either zero or the proportional to the identity map. In other words,

$$\text{if } \forall g \in G, \ \Phi\rho(g) = \sigma(g)\Phi \implies \begin{cases} \Phi \propto \mathbb{I} \text{ if } (\rho, V) = (\sigma, W) \\ \Phi = 0 \text{ if else} \end{cases}$$

Equivalently, if (ρ, V) and (σ, W) are irreducible representations, the space of intertwiners of representations satisfies

$$\mathrm{Hom}_G[(\rho, V), (\sigma, W)] \cong \begin{cases} \mathbb{C} \text{ if } (\rho, V) = (\sigma, W) \\ 0 \text{ if else} \end{cases}$$

A corollary of Schur's lemma is the following: Let (ρ, V) be a irreducible representation of G. Let $M \in \mathbb{C}^{d_\rho \times d_\rho}$ be a matrix. Suppose that

$$\forall g \in G, \quad \rho(g)M = M\rho(g)$$

holds. Then, M is proportional to the identity matrix. The constant of proportionally can be determined by taking traces. Specifically,

$$M = \frac{\mathrm{Tr}[M]}{d_\rho} \mathbb{I}_{d_\rho}$$

Schur's lemma is the key result of representation theory. Schur's lemma

A. Extended Shur Lemma

Schur's Lemma can be extended to reducible representations. Let (ρ, V_ρ) and (σ, V_σ) be G representations which decompose into irridicibles as

$$(\rho, V_\rho) = U[\bigoplus_{\tau \in \hat{G}} m_\tau^\rho (\tau, W_\tau)]U^\dagger \quad (\sigma, V_\sigma) = V[\bigoplus_{\tau \in \hat{G}} m_\tau^\sigma (\tau, W_\tau)]V^\dagger$$

where U, V are fixed unitary matrices that diagonalize the ρ and σ representations, respectively. Then, the vector space of intertwiners between (ρ, V_ρ) and (σ, V_σ) has dimension

$$\dim \mathrm{Hom}_G[(\rho, V_\rho), (\sigma, V_\sigma)] = \sum_{\tau \in \hat{G}} m_\tau^\rho m_\tau^\sigma$$

Furthermore, elements of the space $\mathrm{Hom}_G[(\rho, V_\rho), (\sigma, V_\sigma)]$ have block structure. Specifically, any $\Phi \in \mathrm{Hom}_G[(\rho, V_\rho), (\sigma, V_\sigma)]$ can be parameterized in block diagonal form as

$$\Phi = U[\bigoplus_{\tau \in \hat{G}} \Phi^\tau \otimes \mathbb{I}_{d_\tau}]V^\dagger$$

and each block Φ^τ is a $m_\tau^\rho \times m_\tau^\sigma$ matrix written as

$$\Phi^\tau = \begin{bmatrix} \Phi_{11}^\tau & \Phi_{12}^\tau & \dots & \Phi_{1m_\tau^\sigma}^\tau \\ \Phi_{21}^\tau & \Phi_{22}^\tau & \dots & \Phi_{2m_\tau^\sigma}^\tau \\ \dots & \dots & \dots & \dots \\ \Phi_{m_\tau^\rho 1}^\tau & \Phi_{m_\tau^\rho 2}^\tau & \dots & \Phi_{m_\tau^\rho m_\tau^\sigma}^\tau \end{bmatrix}$$

where each $\Phi_{ij}^\tau \in \mathbb{C}$ is a complex constant and $d_\tau = \dim(\tau, W_\tau)$ is the dimension of the irreducible G-representation (τ, W_τ).

V. INDUCED AND RESTRICTED REPRESENTATIONS OF COMPACT GROUPS

We naturally understand that the group of in-plane rotations is a subgroup of the set of all three dimensional rotations. The Induced and Restricted functors provide a way to generate representations of a subgroup from representations of a larger group and vice-versa. This is especially important in physics, where the number of gapless modes in symmetry breaking can be determined by restriction representations [32]. Similarly, emergent larger symmetries can be understood by induced representations, due to the universality property (ref).

A. Restricted Representation

Let $H \subseteq G$. Let (ρ, V) be a representation of G. The restricted representation of (ρ, V) from G to H is denoted as $\mathrm{Res}_H^G[(\rho, V)]$. Intuitively, $\mathrm{Res}_H^G[(\rho, V)]$ can be viewed as (ρ, V) evaluated on the subgroup H. Specifically,

$$\forall v \in V, \quad \mathrm{Res}_H^G[\rho](h)v = \rho(h)v \tag{3}$$

Note that the restricted representation and the original representation both live on the same vector space V.

B. Induced Representation

The induction representation is a way to construct representations of a larger group G out of representations of a subgroup $H \subseteq G$. Let (ρ, V) be a representation of H. The induced representation of (ρ, V) from H to G is denoted as $\mathrm{Ind}_H^G[(\rho, V)]$. Define the space of functions

$$\mathcal{F} = \{ \ f \ | \ f : G \to V, \ \forall h \in H, \ f(gh) = \rho(h^{-1})f(g) \ \}$$

Then the induced representation is defined as $(\pi, \mathcal{F}) = \mathrm{Ind}_H^G[(\rho, V)]$ where the induced action π acts on the function space $\mathcal{F}$ via

$$\forall g, g' \in G, \ \forall f \in \mathcal{F} \quad (\pi(g) \cdot f)(g') = f(g^{-1}g')$$

1. *Induced Representation for Finite Groups*

There is also an equivalent definition of the induced representation for finite groups that is slightly more intuitive [8]. Let G be a group and let $H \subseteq G$. The set of left cosets of G/H form a partition of G so that

$$G = \bigcup_{i=1}^{|G/H|} g_i H$$

where $\{g_i\}_{i=1}^{|G/H|}$ are a set of representatives of each unique left coset. Note that the choice of left coset representatives is not unique. Now, left multiplication by the element $g \in G$ is an automorphism of G. Left multiplication by $g \in G$ must thus permute left cosets of G/H so that

$$\forall g \in G, \quad g \cdot g_i = g_{j_g(i)} h_i(g)$$

where $j_g : \{1, 2, ..., m\} \to \{1, 2, ..., m\} \in S_m$ is a permutation of left coset representatives. The $h_i(g) \in H$ is an element of subgroup H. The map $j_g(i)$ and group element $h_i(g) \in H$ satisfy a compositionality property. Specifically, we have that

$$\forall g, g' \in G, \quad j_{g'} \circ j_g = j_{g'g}, \quad h_i(g'g) = h_{j_g(i)}(g') \cdot h_i(g)$$

which can be seen by acting on the left cosets with g followed by g' versus acting on the left cosets with $g'g$. Note that

$$e \cdot g_i = g_i \cdot e = g_{j_e(i)} h_i(e)$$

holds so $j_e = e$ and $h_i(e) = e$ holds. Now, let (ρ, V) be a representation of the group H. Let us define the vector space W as

$$W = \bigoplus_{i=1}^{|G/H|} g_i V_{(i)}$$

where the (standard albeit somewhat confusing) notation $g_i V_{(i)}$ denotes an independent copy of the vector space V. This notation is simply a labeling and all copies of $g_i V_{(i)}^H$ are isomorphic to V^H,

$$V \cong g_1 V_1 \cong g_2 V_2 \cong ... \cong g_{|G/H|} V_{|G/H|}$$

so that the space $W \cong \bigoplus_{i=1}^{|G/H|} V$ is just $|G/H|$ independent copies of V. The induced representation lives on this vector space, $(\pi, W) = \mathrm{Ind}_H^G[(\rho, V)]$. The induced action $\pi = \mathrm{Ind}_H^G \rho$ acts on the vector space W via

$$\forall g \in G, \ \forall w = \sum_{i=1}^{|G/H|} g_i v_i \in W, \quad \pi(g) \cdot w = \sum_{i=1}^{|G/H|} \sigma(h_i(g)) v_{j_g(i)} \in W$$

where $v_i \in V_{(i)}$ is in the i-th independent copy of the vector space V. Using the compositionality property of j_g and $h_i(g)$, it is easy to see that this is a valid group action so that $(\pi, W) = \mathrm{Ind}_H^G[(\rho, V)]$ is a valid representation. Note that the induced action π acts on the vector space W by permuting and left action by the H-representation $\rho(h)$.

2. *Universal Property of Induced Representation*

A standard result in group theory establishes the following universal property of induced representations, as stated in [7]:

Theorem 1. *Let $H \subseteq G$. Let (ρ, V) be any H-representation. Let $Ind_H^G(\rho, V)$ be the induced representation of (ρ, V) from H to G. Then, there exists a unique H-equivariant linear map $\Phi_\rho : V \to Ind_H^G V$ such that for any G-representation (σ, W) and any H-equivariant linear map $\Psi : V \to W$, there is a unique G-equivariant map $\Psi^\uparrow : Ind_H^G V \to W$ such that the diagram 3 is commutative.*

$$
\text{L0.3} \qquad
\begin{tikzcd}
(\rho, V) \arrow[r, "\Phi_\rho"] \arrow[dr, "\Psi"'] & \mathrm{Ind}_H^G(\rho, V) \arrow[d, "\Psi^{\uparrow}"] \\
& (\sigma, W)
\end{tikzcd}
$$

FIG. 3: Commutative Diagram for Uniqueness Property of Induced Representations.

Let (ρ, V) be a H-representation and let (σ, W) be a G-representation. Let $\Psi : V \to W$ where Ψ is an intertwiner of a the H-representation and the restriction of the G-representation to an H-representation so that

$$\forall h \in H, \quad \Psi\rho(h) = \mathrm{Res}_H^G[\sigma](h)\Psi$$

so that $\Psi \in \mathrm{Hom}_H[(\rho, V), \mathrm{Res}_H^G(\sigma, W)]$. The universal property of the induced representation allows us to write any such Ψ in a canonical form. Specifically, as illustrated in Figure V B 2, we can always uniquely decompose $\Psi = \Psi^{\uparrow} \circ \Phi_\rho$ where $\Psi^{\uparrow} \in \mathrm{Hom}_G[\mathrm{Ind}_H^G(\rho, V), (\sigma, W)]$ and $\Psi_\rho : V \to \mathrm{Ind}_H^G V$ is (σ, W) independent.

$$
\begin{tikzcd}
(\rho, V) \arrow[r, "\Psi"] \arrow[d, "\rho(h)"'] & (\sigma, W) \arrow[d, "\sigma(h)"', shift right] \arrow[d, "\sigma(g)", shift left] \\
(\rho, V) \arrow[r, "\Psi"] & (\sigma, W)
\end{tikzcd}
\quad \cong \quad
\begin{tikzcd}
(\rho, V) \arrow[rr, "\Psi"] \arrow[dr, "\Phi_\rho"] \arrow[ddd, "\rho(h)"'] & & (\sigma, W) \arrow[ddd, "\sigma(h)"', shift right] \arrow[ddd, "\sigma(g)", shift left] \\
& \mathrm{Ind}_H^G(\rho, V) \arrow[ur, "\Psi^{\uparrow}"] \arrow[d, "{[\mathrm{Ind}_H^G\rho](h)}"', shift right] \arrow[d, "{[\mathrm{Ind}_H^G\sigma](g)}", shift left] & \\
& \mathrm{Ind}_H^G(\rho, V) \arrow[dr, "\Psi^{\uparrow}"] & \\
(\rho, V) \arrow[rr, "\Psi"] \arrow[ur, "\Phi_\rho"] & & (\sigma, W)
\end{tikzcd}
$$

FIG. 4: Factorization Identity for Universal Property of Induced Representations

3. *A Completeness Property For Induced Representations*

Can every function $f : G \to \mathbb{R}^c$ be realized as the induced mapping of functions in $\mathbb{R}^H$? We show that this is the case. We have the following compositional property of induced representations [8]: Let $K \subseteq H \subseteq G$. Let (ρ, V) be any representation of K. Then,

$$\mathrm{Ind}_K^G[(\rho, V)] = \mathrm{Ind}_H^G[\mathrm{Ind}_H^K[(\rho, V)]] \tag{4}$$

which states that the induced representation of (ρ, V) from K to G can be constructed by first inducing (ρ, V) from K to H and then inducing from H to G.

Now, choose $K = \{e\}$ to be the identity element of G. Let (ρ, V) be the trivial one dimensional representation of $K = \{e\}$ with

$$\dim V = 1, \quad \rho(e)v = v$$

Consider the set of left cosets of H in $K = \{e\}$. We have that

$$H/K = H/\{e\} = \{he | h \in G\} = H$$

so the set of coset representatives of H/K is just elements of H. Using a from [8], the induced representation of (ρ, V) from $K = \{e\}$ to H is the left regular representation of H. By the same argument, the induced representation of (ρ, V) from $K = \{e\}$ to G is the left regular representation of G. Thus,

$$\mathrm{Ind}_K^H[(\rho, V)] = (L, \mathbb{C}^H), \quad \mathrm{Ind}_K^G[(\rho, V)] = (L, \mathbb{C}^G)$$

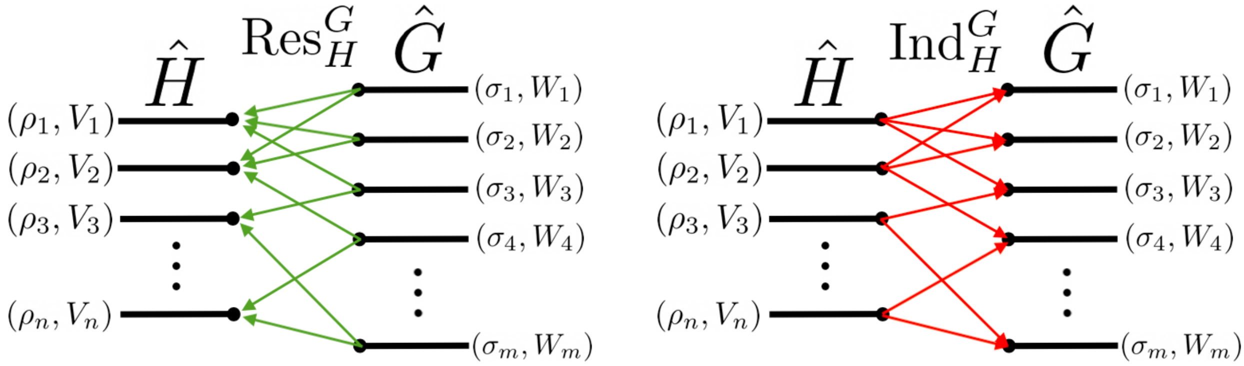

FIG. 5: Left: Restricted representation Res_H^G from G to H of G-irreducibles (σ_i, W_i) to H-irreducibles (ρ_j, V_j). Not every H-representation can be realized as the restriction of a G-representation. Right: Induced representation Ind_H^G from H to G of H-irreducibles (ρ_j, V_j) to G-irreducibles (σ_i, W_i). Not every H-representation can be realized as the induction of a H-representation. The restriction and induction operations are adjoint functors. In general, the restriction and induction operations are generically sparse. This sparsity places restrictions on what irreducibles can appear in $(H \subseteq G)$-equivariant maps.

Using the compositionality property of the induced representation (4), we thus have that

$$(L, \mathbb{C}^G) = \mathrm{Ind}_H^G[(L, \mathbb{C}^H)]$$

Thus, the induced representation from H to G of the left regular representation of H is the left regular representation of G.

$$\begin{tikzcd}
(L, \mathbb{C}^H) \arrow[r, "{\mathrm{Ind}_H^G[(L,\mathbb{C}^H)]}"] \arrow[d, "L(h)"'] & (L, \mathbb{C}^G) \arrow[d, "L(h)"', "L(g)"] \\
(L, \mathbb{C}^H) \arrow[r, "{\mathrm{Ind}_H^G[(L,\mathbb{C}^H)]}"] & (L, \mathbb{C}^G)
\end{tikzcd}$$

FIG. 6: Commutative Diagram for Completeness Property of Induced Representations. L_h denotes the left regular action of H on $\mathbb{C}^H$. L_g denotes the left regular action of G on $\mathbb{C}^G$. The induced representation of the left regular representation of H is the left regular representation of G, $(L, \mathbb{C}^G) = \mathrm{Ind}_H^G[(L, \mathbb{C}^H)]$. The induced representation makes the diagram commutative. This should be contrasted with the definition of G-equivarience defined in II A.

Thus, the induction operation maps the space of all group valued functions on H into the space of all group valued functions on G.

C. $(H \subseteq G)$-Intertwiners

We will also consider another definition of intertwiners between different groups. Let $H \subseteq G$. Let (ρ, V) be a H-representation. Let (σ, W) be a G-representation. We define the vector space of intertwiners of (ρ, V) and (σ, W) as

$$\mathrm{Hom}_H[(\rho, V), \mathrm{Res}_H^G[(\sigma, W)]] = \{\ \Phi \mid \Phi : V \to W, \text{ s.t. } \quad \forall h \in H, \ \ \Phi(\rho(h)v) = \sigma(h)\Phi(v)\ \}$$

We say that a linear map $\Phi : V \to W$ is an $(H \subseteq G)$-intertwiner of the H-representation (ρ, V) and the G-representation (σ, W) if $\Phi \in \mathrm{Hom}_H[(\rho, V), \mathrm{Res}_H^G[(\sigma, W)]]$. The induction and restriction operations are adjoint functors [7]. By the Frobinous reciprocity theorem [7],

$$\mathrm{Hom}_H[(\rho, V), \mathrm{Res}_H^G[(\sigma, W)]] \cong \mathrm{Hom}_G[\mathrm{Ind}_H^G[(\rho, V)], (\sigma, W)]$$

and so for every $\Phi : V \to W$ which intertwines (ρ, V) and $\mathrm{Res}_H^G[(\sigma, W)]$ over H there is a unique $\Phi^{\uparrow} : \mathrm{Ind}_H^G[V] \to W$ that intertwines $\mathrm{Ind}_H^G[(\rho, V)]$ and (σ, W) over G. Not every H-representation can be realized as the restriction of a G-representation. Thus, the universe of $(H \subseteq G)$-intertwiners is a proper subset of the universe of H-intertwiners.

$$\begin{tikzcd}
(\rho, V) \arrow[r, "\Phi"] \arrow[d, "\rho(h)"'] & (\sigma, W) \arrow[d, "\sigma(h)"', "\sigma(g)"] \\
(\rho, V) \arrow[r, "\Phi"] & (\sigma, W)
\end{tikzcd}$$

FIG. 7: Commutative Diagram For $(H \subseteq G)$-intertwiner. $\Phi : V \to W$. The map $\Phi \in \mathrm{Hom}_H[(\rho, V), \mathrm{Res}_H^G[(\sigma, W)]] \cong \mathrm{Hom}_G[\mathrm{Ind}_H^G[(\rho, V)], (\sigma, W)]$ if and only if the following diagram is commutative for all $h \in H$. Note that the group G also has $\sigma(g)$ action on the vector space W.

A map $\Phi : V \to W$ is a $(H \subseteq G)$-intertwiner if and only if the diagram in V C is commutative.

D. Irridducibility and Induced and Restricted Representations

Let H be a subgroup of compact group G. We can use the induced representation to map representations of H to representations of G and the restricted representation to map representations of G to representations of H. All representations of H break down into direct sums of irreducible representations of H. Similarly, all representations of G break down into direct sums of irreducible representations of G. Let use denote $\hat{H}$ as a set of representatives of all irreducible representations of H and $\hat{G}$ as a set of representatives of all irreducible representations of G,

$$\hat{H} = \{ \ (\rho, V_\rho) \ | \ \text{ Representative irreducibles of } H \ \}$$
$$\hat{G} = \{ \ (\sigma, W_\sigma) \ | \ \text{ Representative irreducibles of } G \ \}$$

We want to understand how the restriction and induction operations transform H-irreducibles to G-irreducibles and vice versa. We can completely characterize how irreducibles change under the restriction and induction procedures using *branching rules* and *induction rules*, respectively.

1. Restricted Representation and Branching Rules

Let (σ, W) and (σ', W') be G-representations. The restriction operation is linear and

$$\text{Res}_H^G[(\sigma, W) \oplus (\sigma', W')] = \text{Res}_H^G[(\sigma, W)] \oplus \text{Res}_H^G[(\sigma', W')]$$

We can study the restriction operation by looking at restrictions of the set of G-irreducibles $\hat{G}$. The restriction of an G-irreducible is not necessarily irreducible in H and will decompose as a direct sum of H-irreducibles. Let $(\sigma, W_\sigma) \in \hat{G}$. We can define a set of integers $B_{\sigma,\rho} : \hat{G} \times \hat{H} \to \mathbb{Z}^{\geq 0}$,

$$\text{Res}_H^G[(\sigma, W_\sigma)] = \bigoplus_{\rho \in \hat{H}} B_{\sigma,\rho}(\rho, W_\rho)$$

so that $B_{\sigma,\rho}$ counts the multiplicities of the H-irreducible (ρ, W_ρ) in the restricted representation of the G-irreducible (σ, W_σ). The $B_{\sigma,\rho}$ are called *branching rules* and they have been well studied in the context of particle physics [31]. Let (σ', W') be any G-representation. (σ', W') will decompose into G-irreducibles as

$$(\sigma', W') = \bigoplus_{\sigma \in \hat{G}} m_\sigma(\sigma, W_\sigma)$$

where m_σ counts the number of copies of the G-irreducible (σ, W_σ) in (σ', W'). Then, the restriced representation of (σ', W') decomposes into H-irreducibles as

$$\text{Res}_H^G[(\sigma', W')] = \bigoplus_{\sigma \in \hat{G}} m_\sigma \, \text{Res}_H^G[(\sigma, W_\sigma)] = \bigoplus_{\rho \in \hat{G}} \sum_{\sigma \in \hat{G}} [m_\sigma B_{\sigma,\rho}](\rho, W_\rho)$$

So that the multiplicity of the (ρ, W_ρ) irreducible in the restriction of (σ', W') is $\sum_{\sigma \in \hat{G}} m_\sigma B_{\sigma,\rho}$. Thus, the branching rules $B_{\sigma,\rho}$ completely determine how an arbitrary G-representation restricts to an H-representation.

2. Induced Representation and Induction Rules

The induction operation acts linearly on representations composed of direct sums of representations. Specifically, if (ρ_1, V_1) and (ρ_2, V_2) are representations of H, then

$$\text{Ind}_H^G[(\rho_1, V_1) \oplus (\rho_2, V_2)] = \text{Ind}_H^G[(\rho_1, V_1)] \oplus \text{Ind}_H^G[(\rho_2, V_2)]$$

The induction operation Ind_H^G maps every irreducible representation $(\rho, V_\rho) \in \hat{H}$ to a G-representation. The induced representation of an irreducible representation of H is not necessarily irreducible in G and will break into irreducibles in $\hat{G}$ as

$$\text{Ind}_H^G[(\rho, V_\rho)] = \bigoplus_{\sigma \in \hat{G}} I_{\rho,\sigma}(\sigma, W_\sigma)$$

where the integers $I_{\rho,\sigma} : \hat{H} \times \hat{G} \to \in \mathbb{Z}^{\geq 0}$ denotes the number of copies of the irreducible $(\sigma, W_\sigma) \in \hat{G}$ in the induced representation $\text{Ind}_H^G(\rho, V_\rho)$ of the irreducible (ρ, V_ρ). The $I_{\rho,\sigma}$ are called *Induction Rules* and completely determine the multiplicities of G-irreducibles in the induced representation of any H-representation. Specifically, let (ρ', V') be any representation of H. Then, (ρ', V') breaks into H-irreducibles as

$$(\rho', V') = \bigoplus_{\rho \in \hat{H}} n_\rho (\rho, V_\rho)$$

The induced representation is linear and maps (ρ', V') into a representation of G which will break into G-irreducibles as

$$\mathrm{Ind}_H^G[(\rho', V')] = \bigoplus_{\rho \in \hat{H}} n_\rho \,\mathrm{Ind}_H^G(\rho, V_\rho) = \bigoplus_{\sigma \in \hat{G}} (\sum_{\rho \in \hat{H}} n_\rho I_{\rho,\sigma})(\sigma, W_\sigma)$$

so that the multiplicity of $(\sigma, W_\sigma) \in \hat{G}$ in the induced representation of $(\rho, V_\rho) \in \hat{H}$ is given by $\sum_{\rho \in \hat{H}} m_\sigma I_{\rho,\sigma}$. Thus, the induction rules $I_{\rho,\sigma}$ completely determine the multiplicities of G-representations in the induced representation of any H-representation.

3. Irreducibility and Frobenius Reciprocity

The induction rules $I_{\rho\sigma} : \hat{H} \times \hat{G} \to \mathbb{Z}^{\geq 0}$ and the branching rules $B_{\sigma\rho} : \hat{G} \times \hat{H} \to \mathbb{Z}^{\geq 0}$ are related by the Frobenius reciprocity theorem [7]. Let (ρ', V') be any H-representation and let (σ', W') be any G-representation. Then,

$$\mathrm{Hom}_H[(\rho', V'), \mathrm{Res}_H^G[(\sigma', W')]] \cong \mathrm{Hom}_G[\mathrm{Ind}_H^G[(\rho', V')], (\sigma', W')]$$

Choosing $(\rho', V') = (\rho, V_\rho) \in \hat{H}$ and $(\sigma', W') = (\sigma, W_\sigma) \in \hat{G}$ gives $I_{\rho,\sigma} = B_{\sigma,\rho}$. So that when viewed as matrices, $B = I^T$. All information about how H-representations are induced to G-representations and G-representations are restricted to H-representations is encoded in both $B_{\sigma,\rho}$ and $I_{\rho,\sigma}$. It should be noted for many cases of interest, $B_{\sigma,\rho}$ and $I_{\rho,\sigma}$ are sparse, and have non-zero entries for only a small number of ρ and σ pairs.

VI. TENSOR PRODUCT REPRESENTATIONS

The tensor product of two group representations is itself a group representation. Given two representations (ρ, V) and (σ, W) of a compact group G, their tensor product is defined as a new representation where the group action on $v \otimes w \in V \otimes W$ is defined via:

$$(\rho \otimes g)(g) \cdot (v \otimes w) = (\rho(g) \cdot v) \otimes (\sigma(g) \cdot w), \quad \text{for all } g \in G,\ v \in V,\ w \in W.$$

This construction is a systematic way to combine two representations into a single, higher-dimensional representation.

A. Decomposition into Irreducibles

A key property of compact groups is that their representations are fully reducible, meaning any representation, including a tensor product representation, can be expressed as a direct sum of irreducible representations. Thus, the tensor product of two representations can always be decomposed as:

$$\rho \otimes \sigma = U_{\rho\sigma} [\bigoplus_{\tau \in \hat{G}} m_{\rho\sigma}^\tau \tau] U_{\rho\sigma}^\dagger,$$

where $U_{\rho\sigma}$ is a unitary matrix and $m_{\rho\sigma}^\tau$ are a set of integers that describe how many copies of the irreducible representation τ appear in the tensor product $\rho \otimes \sigma$. This decomposition is central to many applications, as it reveals how the combined system transforms in terms of the simpler irreducible components. The matrix $U_{\rho\sigma}$, which describe the change of basis, are sometimes referred to as Clebsch-Gordon coefficients. These coefficients arise naturally in the decomposition of tensor product representations, particularly in the context of rotation groups such as $SU(2)$ or $SO(3)$, which are widely used in quantum mechanics.

1. *Example: $SU(2)$ Tensor Product Representations*

These coefficients are essential in angular momentum coupling, such as combining the spins of particles or the orbital and spin angular momentum of a single particle. The coefficients appear in spherical harmonics and are used in problems involving symmetry and wave functions on the sphere. For the group $SU(2)$, irreducible representations are labeled by their "spin" j, which can take values $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. The tensor product of two irreps labeled by j_1 and j_2 decomposes as:

$$D^{(j_1)} \otimes D^{(j_2)} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} D^{(j)},$$

where $D^{(j)}$ is the irreducible representation of spin j. Clebsch-Gordon coefficients describe the change of basis between two natural bases. The basis of the tensor product representation $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$, where m_1 and m_2 are the magnetic quantum numbers. The basis of the irreducible components $|j, m\rangle$, where $m = m_1 + m_2$ and j is the total angular momentum. The relationship is expressed as:

$$|j, m\rangle = \sum_{m_1, m_2} C^{j,m}_{j_1,m_1;j_2,m_2} |j_1, m_1\rangle \otimes |j_2, m_2\rangle,$$

where $C^{j,m}_{j_1,m_1;j_2,m_2}$ are the Clebsch-Gordon coefficients.

VII. IRREDUCIBLE REPRESENTATIONS OF THE SYMMETRIC GROUP

The symmetric group on n elements, denoted as S_n, is the set of bijections from the set $\{1, 2, ..., n\}$ into itself. The size of the symmetric group $|S_n| = n! \sim \exp(-n)n^n$ which grows super-exponentially in n. The fundamental (or matrix) representation of S_n is the $n \times n$ representation

$$F(\sigma)_{ij} = \begin{cases} 1 \text{ if } i = \sigma(j) \\ 0 \text{ else} \end{cases}$$

The fundamental representation is reducible. To see this, note that the subspace spanned by the sum of the Euclidean basis vectors is an invariant subspace. Specifically, we have that

$$\forall \sigma \in S_n, \quad F(\sigma)[\sum_{i=1}^{n} e_i] = \sum_{i=1}^{n} e_{\sigma(i)} = \sum_{i=1}^{n} e_i$$

Irreducible Representations of S_n are particularly elegant [6]. A classic result in group theory states that irreducible representations of S_n are indexed by partitions of n. Specifically, for every partition $\lambda \vdash n$ there is a unique irreducible representation of S_n. The structure of irreducible representations of S_n can be understood with the help of Specht Modules. In order to work with Specht Modules, we introduce Young diagrams and Young tableau.

A. Young Diagrams and Young Tableau

Young diagrams are combinatorial tools used in the study of symmetric groups. They provide a visual way to describe partitions of integers and play a central role in understanding irreducible representations.

A **Young diagram** is a collection of boxes arranged in left-aligned rows, where the number of boxes in each row corresponds to the parts of a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ of a positive integer n. A partition λ satisfies:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0, \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = n.$$

Here, $\lambda \vdash n$ indicates that λ is a partition of n. For example, the partition $\lambda = (4, 2, 1)$ corresponds to the Young diagram:

$$\begin{array}{llll} \square & \square & \square & \square \\ \square & \square & & \\ \square & & & \end{array}$$

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline 7 \\ \cline{1-1} 8 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 3 & 4 & 5 \\ \hline 6 & 7 & 8 \\ \cline{1-3} 9 & 10 \\ \cline{1-2}\end{array}$$

FIG. 8: Canonical Young Tableau: The partition $\lambda = (4, 2, 2) \vdash 8$. The associated canonical Young Tableau $\hat{\lambda}$ is shown in

For the symmetric group S_n, irreducible representations correspond to partitions of n. Young diagrams provide a convenient way to label and study these representations. Young diagrams help describe the branching rules when restricting representations of S_n to S_{n-1}. A **standard Young tableau** is a Young diagram where the boxes are filled with integers $1, 2, \ldots, n$, such that the numbers increase along each row and the numbers increase down each column. For example, a standard Young tableau for $\lambda = (3, 2)$ is:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & \end{array}$$

The number of standard Young tableaux of shape λ is given by the *hook-length formula* [18], which is central in the representation theory of symmetric groups.

B. Specht Modules

Specht modules provide a general method for constructing the irreducible representations of S_n. Let $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m) \vdash n$ be a partition of n with $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_m$. To each permutation $\lambda| \rightarrow \hat{\lambda}$, we associate a Young diagram,

We then define the following subgroups of S_n. We have that

$$P_\lambda = \{g \in S_n | \text{g preserves the rows of the } \lambda\text{-tableau}\}$$
$$Q_\lambda = \{g \in S_n | \text{g preserves the columns of the } \lambda\text{-tableau}\}$$

To each of these subgroups we defined the group algebra elements

$$p_\lambda = \sum_{g \in P_\lambda} g, \quad q_\lambda = \sum_{g \in Q_\lambda} \text{sign}(g) g$$

The sub-algebra $V_\lambda = \mathbb{C}[S_n] p_\lambda q_\lambda$ forms an irreducible representation of S_n. The dimension of an irreducible of S_n can be calculated using the *hook length formula*, [18]. Specifically, let $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m) \vdash n$.

C. Hook Length Formula

The dimension of an irreducible representation of S_n corresponding to a partition λ can be calculated using the *hook length formula* [18]. Specifically, let

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \vdash n,$$

where λ is a partition of n, and λ_i are the parts of the partition, satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ and $\sum_{i=1}^m \lambda_i = n$.

To compute the dimension of the irreducible representation V_λ, construct the *Young diagram* of the partition λ. The Young diagram is a collection of boxes arranged in rows, where the i-th row has λ_i boxes. For each box in the Young diagram, its *hook length* is defined as the number of boxes directly to the right, directly below, or in the same position, including the box itself. Denote the hook length of a box (i, j) in the diagram as $h(i, j)$. The dimension of V_λ is then given by:

$$\dim(V_\lambda) = \frac{n!}{\prod_{(i,j) \in \lambda} h(i, j)},$$

where the product runs over all boxes (i, j) in the Young diagram of λ.

D. Example: Symmetric Group S_4

Consider the symmetric group S_4. One possible partition of 4 is $\lambda = (2, 2)$, which corresponds to the Young diagram:

$$\begin{array}{cc} \square & \square \\ \square & \square \end{array}$$

The hook lengths for this diagram are:

$$\begin{array}{cc} 3 & 2 \\ 2 & 1 \end{array}$$

Using the formula:

$$\dim(V_\lambda) = \frac{4!}{3 \cdot 2 \cdot 2 \cdot 1} = \frac{24}{12} = 2.$$

This means the irreducible representation corresponding to $\lambda = (2, 2)$ has dimension 2.

E. Pieri's Formula in Representation Theory of Compact Groups

Pieri's formula describes how the tensor product of an irreducible representation with a fundamental representation decomposes into a direct sum of irreducible representations. While originally developed in the context of symmetric groups, it generalizes to compact groups, particularly $GL(n)$, in terms of highest weights and partitions.

1. *Statement of Pieri's Formula*

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ represent the highest weight of an irreducible representation V_λ of $GL(n)$, corresponding to the Young diagram of the partition λ. Let $\mu = (1)$, the fundamental (vector) representation. Then, the tensor product of V_λ with V_μ decomposes as:

$$V_\lambda \otimes V_\mu = \bigoplus_\nu V_\nu,$$

where the sum is over all partitions ν obtained by adding one box to the Young diagram of λ, subject to the condition that the resulting diagram is still a valid Young diagram.

2. *Illustrative Example*

Consider $\lambda = (2, 1)$, represented by the Young diagram:

$$\begin{array}{cc} \square & \square \\ \square & \end{array}$$

When $\mu = (1)$, the partitions ν resulting from adding one box are:

$$(3, 1), \quad (2, 2), \quad (2, 1, 1).$$

Thus:

$$V_{(2,1)} \otimes V_{(1)} = V_{(3,1)} \oplus V_{(2,2)} \oplus V_{(2,1,1)}.$$

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \otimes \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 6 & 7 \\ \hline\end{array}$$

FIG. 9: Diagrammatic Computation of Tensor Products of Young Tableau: The partition $\lambda = (4,2,2) \vdash 8$. The associated canonical Young Tableau $\hat{\lambda}$ is shown in

3. Tenor Products of Irreducible Representations of S_n

Let λ and λ' be two irreducible representations of S_n, the tensor product

$$V_\lambda \otimes V_{\lambda'} = U_{\lambda\lambda'}[\sum_{\tau \vdash n} c^\tau_{\lambda\lambda'} V_\tau] U^\dagger_{\lambda\lambda'}$$

where $U_{\lambda\lambda'}$ is a unitary change of basis. The integers $c^\tau_{\lambda\lambda'}$ are called the Littlewood-Richardson coefficients and count the number of irreducible τ-representations in the tensor product of λ and λ'. The tensor product coefficients $c^\tau_{\lambda\lambda'}$ are generically sparse and the relation

$$C^\tau_{\lambda\lambda'} > 0 \implies |d_\lambda - d_{\lambda'}| \leq d_\tau \leq d_\lambda + d_{\lambda'}$$

holds [18]. Tensor product decomposition's can be computed diagrammatically using the Littlewood-Richerdson rule. Consider the tensor product decomposition of two

F. Irreducible Decomposition of Tensor Product of N Identical Vector Spaces

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. In many linear algebra applications, we often work with a vector space W that is composed of k-fold tensor products of the smaller vector space V such that

$$W = V^{\otimes k} = \underbrace{V \otimes V \otimes ... \otimes V}_{k-\text{times}}$$

where the dimension of the vector space V is d, $\dim V = d$ so that the dimension of W is $\dim W = d^k$. This situation arises naturally in dealing with quantum mechanical systems of many identical particles. Tensor product representations of this form are naturally related to the permutation group. Specifically, let $W = V^{\otimes k}$ be a vector space that is the k-fold tensor product of V. For each permutation $\sigma \in S_k$, we define the operator $\hat{S}_\sigma$ with action on the tensor product basis via permutation

$$\forall \sigma \in S_k, \quad \hat{S}_\sigma |i_1 i_2 ... i_k\rangle = |i_{\sigma(1)} i_{\sigma(2)} ... i_{\sigma(k)}\rangle$$

The operators $\hat{S}_\sigma$ form a unitary reducible representation of the group S_n. Specifically, the permutation representation will decompose as

$$(\hat{S}_\sigma, V^{\otimes k}) \cong \bigoplus_{\lambda \vdash k} c^\lambda_{(k,d)} \lambda$$

with $c^\lambda_{(k,d)}$ counting the muplicity of the irreducible λ representation in $(\hat{S}_\sigma, V^{\otimes k})$. The character of the $(\hat{S}_\sigma, V^{\otimes k})$ representation is given by

$$\chi(\sigma) = \mathrm{Tr}[\hat{S}_\sigma] = d^{f(\sigma)}$$

where $f(\sigma)$ is the number of fixed points of the permutation σ. Thus,

$$c^\lambda_{(k,d)} = \sum_{\sigma \in S_k} \chi_\lambda(\sigma) d^{f(\sigma)}$$

where $\chi_\lambda(\sigma) : S_k \to \mathbb{C}$ is the character of the λ irreducible. Suppose that the matrix X commutes with all N-fold products of unitary matrices

$$\forall U \in G, \ \ U^{\otimes N} X = X U^{\otimes N}$$

holds, where G is either $O(d)$ or $U(d)$. Then the matrix X may be written as

$$X = \sum_{\sigma \in S_k} c_\sigma \hat{S}_\sigma$$

for some complex constants c_σ. There are $N!$ constants c_σ, one for each permutation σ in S_N. A set of $N!$ linear equations for the coefficients c_σ can be computed by multiplication by an element $\hat{S}_\tau, \tau \in S_N$ and taking traces [22].

1. *Example: $N = 2$ Case*

For $N = 2$, $S_2 \cong \mathbb{Z}_2$ is isomorphic to the cyclic group of order two. There are two permutation operators, $\mathbb{1}$ and $\hat{S}$. The operator $\hat{S}$ permutes tensor product indices with $\hat{S}|ij\rangle = |ji\rangle$. Note that

$$\hat{S}^2 = \mathbb{1}$$

Thus, $\hat{S}$ has eigenvalues ± 1. All representations of S_2 are one dimensional. There are two irreducible representations, the trivial and sign representation. The tensor product space then decomposes as

$$V \otimes V = [\frac{d(d+1)}{2} V_+] \bigoplus [\frac{d(d-1)}{2}] V_-$$

so that the tensor permutation space decomposes into $\frac{d(d+1)}{2}$ copies of the symmetric space and $\frac{d(d-1)}{2}$ copies of the anti-symmetric space. The projection operators into the V_+ and V_- subspaces are given by

$$\hat{S}_+ = \frac{1}{\sqrt{2}}(\mathbb{1}_{d\times d} + \hat{S}) \quad \hat{S}_- = \frac{1}{\sqrt{2}}(\mathbb{1}_{d\times d} - \hat{S})$$

respectively. The projection operators are normalized to satisfy the relations $\hat{S}_\pm^2 = \hat{S}_\pm$. Using Young diagrams, the irreducible representations are representation as the partitions $\lambda \vdash 2$, as shown in 10.

$$V_+ \cong V_{(2)} \cong \boxed{1}\boxed{2}, \quad V_- \cong V_{(1,1)} \cong \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array}$$

FIG. 10: Irreducible Representations of S_2 and corresponding Young Diagrams

2. *Example: Tensor Product Rules of S_2*

The tensor product rules for the group S_2 are trivial. Using characters, we have that

$$V_+ \otimes V_+ = V_+, \quad V_+ \otimes V_- = V_-, \quad V_- \otimes V_- = V_+$$

so that $C_+^{++} = 1,\ C_-^{+-} = C_-^{+-} = 1,\ C_+^{--} = 1$ and all other tensor product multiplicities are zero.

3. *Example: Computing Branching and Induction Rules of $S_2 \times S_2 \subseteq S_4$*

There are five irreducible representations of S_4. The character table of irriducibles of S_4. The group S_4 has five conjugacy classes.
Evaluated on the $S_2 \times S_2$ subgroup, we have that

$$\begin{aligned}
&\chi_{(4)}[(e)(e)] = 1, \quad \chi_{(4)}[(12)(e)] = 1, \quad \chi_{(4)}[(e)(34)] = 1, \quad \chi_{(4)}[(12)(34)] = 1\\
&\chi_{(1,1,1,1)}[(e)(e)] = 1, \quad \chi_{(1,1,1,1)}[(12)(e)] = -1, \quad \chi_{(1,1,1,1)}[(e)(34)] = -1, \quad \chi_{(1,1,1,1)}[(12)(34)] = 1\\
&\chi_{(2,2)}[(e)(e)] = 2, \quad \chi_{(2,2)}[(12)(e)] = 0, \quad \chi_{(2,2)}[(e)(12)] = 0, \quad \chi_{(2,2)}[(12)(12)] = 2\\
&\chi_{(2,1,1)}[(e)(e)] = 3, \quad \chi_{(2,1,1)}[(12)(e)] = 1, \quad \chi_{(2,1,1)}[(e)(12)] = 1, \quad \chi_{(2,1,1)}[(12)(12)] = -1,\\
&\chi_{(3,1)}[(e)(e)] = 3, \quad \chi_{(3,1)}[(12)(e)] = -1, \quad \chi_{(3,1)}[(e)(12)] = -1, \quad \chi_{(3,1)}[(12)(12)] = -1,
\end{aligned}$$

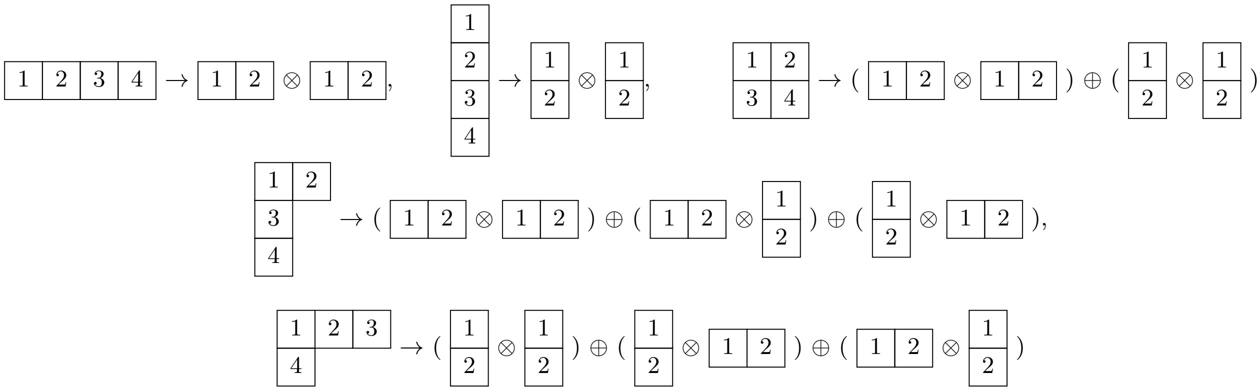

FIG. 11: Under the group restriction operation of S_4 to $S_2 \times S_2$, The five irreducible representations $\lambda \vdash 4$ of S_4 decompose into direct sums of tensor products of S_2 irreducible representations.

Character Table of Irreducible Representations of S_4					
Character	e,(size=1)	(12),(size=6)	(12)(34),(size=3)	(123),(size=8)	(1234),(size=6)
$\chi_{(4)}$	1	1	1	1	1
$\chi_{(1,1,1,1)}$	1	-1	1	-1	1
$\chi_{(2,2)}$	2	0	2	-1	0
$\chi_{(2,1,1)}$	3	1	-1	0	-1
$\chi_{(3,1)}$	3	-1	-1	0	1

TABLE I: Character Table of S_4 for irreducible representations $\lambda \vdash 4$.

Upon restriction to the subgroup $S_2 \times S_2$ we have the following decomposition of S_4 irreducible representations,

$$\begin{aligned}
&V_{(4)} \to V_+ \otimes V_+, \quad V_{(1,1,1,1)} \to V_- \otimes V_-, , \quad V_{(2,2)} \to (V_+ \otimes V_+) \oplus (V_- \otimes V_-) \\
&V_{(2,1,1)} \to (V_+ \otimes V_+) \oplus (V_+ \otimes V_-) \oplus (V_- \otimes V_+), \quad V_{(3,1)} \to (V_- \otimes V_-) \oplus (V_+ \otimes V_-) \oplus (V_- \otimes V_+)
\end{aligned}$$

This is shown diagrammatically in 12. Thus, the only non-zero branching rules are given by

$$\begin{aligned}
&B^{++}_{(4)} = 1, \quad B^{--}_{(1,1,1,1)} = 1, \quad B^{++}_{(2,2)} = B^{--}_{(2,2)} = 1 \\
&B^{--}_{(3,1)} = B^{+-}_{(3,1)} = B^{-+}_{(3,1)} = 1, \quad B^{++}_{(2,1,1,1)} = B^{+-}_{(2,1,1)} = B^{-+}_{(2,1,1)} = 1
\end{aligned}$$

VIII. SCHUR-WEYL DUALITY

Schur-Weyl Duality is a powerful tool in the representation theory of compact groups [27]. In the literature there is some ambiguity as to the actual definition of what Schur-Weyl duality entails. Schur-Weyl Duality is sometimes referred to as the decomposition of the tensor products classical Lie groups. However, Schur-Weyl is actually a more general idea that can be used to decompose any k-fold tensor product of a representation of a compact group. Specifically, the k-fold tensor product of a representation of a compact group G forms a representation of both the group G and the group $G \times S_k$. Using the representation theory of the symmetric group VII, we can decompose the k-fold tensor product into representation of G and representations of S_k.

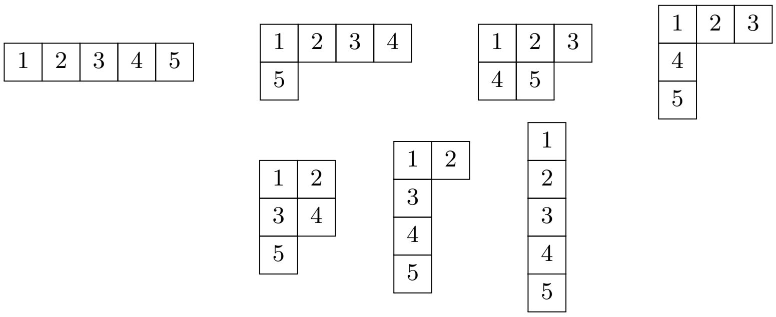

FIG. 12: The seven integer partitions of five. Each partition $\lambda \vdash 5$ is in bijective correspondence with a irreducible representation of S_5. The dimensions of the corresponding irreducible representation λ going from left to right are: 1, 1,4,4,5,5,6

A. General Schur-Weyl Duality

Let G be a compact group. Let (ρ, V_ρ) be any representation of G. Consider the k-fold tensor product representation, $(\rho^{\otimes k}, V_\rho^{\otimes k})$. This representation also forms a representation of the symmetric group of order k, as

$$\forall \sigma \in S_k,\ \ \forall g \in G, \quad S_\sigma \underbrace{[\rho(g) \otimes \rho(g) \otimes ... \otimes \rho(g)]}_{k-times} = \underbrace{[\rho(g) \otimes \rho(g) \otimes ... \otimes \rho(g)]}_{k-times} S_\sigma$$

so that the G action and S_k action are commutative.

$$\begin{tikzcd} V_\rho^{\otimes k} \arrow[r, "\rho^{\otimes k}(g)"] \arrow[d, "\hat{S}_\sigma"'] & V_\rho^{\otimes k} \arrow[d, "\hat{S}_\sigma"'] \\ V_\rho^{\otimes k} \arrow[r, "\rho^{\otimes k}(g)"] & V_\rho^{\otimes k} \end{tikzcd}$$

FIG. 13: 'Square'-type commutative diagram for Schur-Weyl duality. The key observation in Schur-Weyl duality is that the k-fold tensor product action and the tensor permutation representation are commutative. This allows for definition of $G \times S_k$ action on the vector space $V_\rho^{\otimes k}$. Because of this, $(\Pi_\rho^k, V_\rho^{\otimes k})$ forms a representation of the group $G \times S_k$.

Let us define the action Π_ρ^k on the vector space $V_\rho^{\otimes k}$ as the following

$$\forall g \in G,\ \ \forall \sigma \in S_k,\ \ \forall w_{i_1 i_2 ... i_k} \in V_\rho^{\otimes k}, \quad \Pi_\rho^k(g, \sigma) w_{i_1 i_2 ... i_k} = \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} ... \sum_{j_k=1}^{d} \rho(g)_{i_{\sigma(1)} j_1} \rho(g)_{i_{\sigma(2)} j_2} ... \rho(g)_{i_{\sigma(k)} j_k} w_{j_1 j_2 ... j_k}$$

Note that this action is well defined and can be performed by matrix multiplication followed by permutation or permutation followed by matrix multiplication. For this reason, $(\Pi_\rho^k, V^{\otimes k})$ is a well defined representation of the group $G \times S_k$. The representation $(\Pi_\rho^k, V_\rho^{\otimes k})$ is in general not reducible and will decompose into irreducible representations of $G \times S_k$. Irreducible representations of $G \times S_k$ are tensor products of irreducible representations of G and irreducible representations of S_k. Thus, we have the following decomposition,

$$(\Pi_\rho^k, V_\rho^{\otimes k}) \cong \bigoplus_{\tau \in \hat{G}} \bigoplus_{\lambda \vdash k} m_\rho^{k\tau\lambda} (\tau, V_\tau) \otimes (\lambda, V_\lambda)$$

where $m_\rho^{k\tau\lambda}$ are integers counting the number of copies of the $(\tau, V_\tau) \otimes (\lambda, V_\lambda)$ irreducible in $(\Pi_\rho^k, V_\rho^{\otimes k})$. Thus, the tensor product space decomposes into vector subspaces that are characterized by their transformation properties based on G action and tensor index permutations.

B. Unitary Schur-Weyl Duality

Let us apply the more general Schur-Weyl formalism to the case of the unitary group $U(d)$. Irreducible representations of $U(d)$ are countably infinite and are in one-to-one correspondence with integer partitions [27, 31]. Let

$\lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ be a partition with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m$. The irreducible representation of $U(d)$ associated to the partition λ will be denoted as (U_λ, V_λ). Let $(U_1, \mathbb{C}^d)$ be the fundamental d-dimensional representation of $U(d)$ defined as the $\lambda = (1)$ partition,

$$U_d = \{ \ U \ \mid \ U^\dagger U = \mathbb{I}_d = UU^\dagger \ \}$$

Consider the k-fold tensor product decomposition,

$$(\mathbb{C}^d)^{\otimes k} = \bigoplus_{\lambda \vdash (k,d)} V_\lambda \otimes \lambda$$

where $\lambda \vdash (k,d)$ denotes partitions of the integer k with no more than d summands, i.e.

$$\lambda \vdash (k,d) \implies \lambda = (\lambda_1, \lambda_2, ..., \lambda_m), \quad \text{s.t } \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m \text{ s.t. } \sum \lambda_i = k \text{ and } m \leq d$$

A celebrated theorem of Weyl [27] states that the representations (U_λ, V_λ) exhaust all representations of the d-dimensional unitary group $U(d)$.

1. Unitary Group Tensor Product Rules

For a complete discussion of diagrammatic methods for computing tensor products of irreducible representations of the unitary group, please see [11]. We will be interested in tensor products of irreducible representations of $U(d)$. Let λ and λ' be two partitions. Let V_λ and $V_{\lambda'}$ be the corresponding irreducible representations of $U(d)$. Then, consider the tensor product

$$V_\lambda \otimes V_{\lambda'} \cong \bigoplus_{n=1}^{\infty} \bigoplus_{\mu \vdash n} m^{\mu}_{\lambda\lambda'} V_\mu$$

so that the index μ ranges over all integer partitions and $m^{\mu}_{\lambda\lambda'}$ are integers that count the muplicity of the irreducible representation V_μ in the tensor product $V_\lambda \otimes V_{\lambda'}$. Using Schur-Weyl Duality, we can derive an exact expression for tensor product rules $m^{\mu}_{\lambda\lambda'}$ of the unitary group in terms of the branching rules of the symmetric group. To begin, consider the trivial relation

$$(\mathbb{C}^d)^{\otimes k} \otimes (\mathbb{C}^d)^{\otimes k'} = (\mathbb{C}^d)^{\otimes (k+k')}$$

for any integers k and k'. Then, using the vector space decomposition in the Schur-Weyl duality, we have an isomorphism of vector spaces

$$\underbrace{[\bigoplus_{\lambda \vdash (k,d)} V_\lambda \otimes \lambda]}_{(\mathbb{C}^d)^k} \otimes \underbrace{[\bigoplus_{\lambda' \vdash (k',d)} V_{\lambda'} \otimes \lambda']}_{(\mathbb{C}^d)^{k'}} \cong \underbrace{[\bigoplus_{\mu \vdash (k+k',d)} V_\mu \otimes \mu]}_{(\mathbb{C}^d)^{k+k'}}$$

This is a representation of the group $U(d) \times S_k \times S_{k'}$. Expanding out the tensor product of the left hand side, we have that

$$\bigoplus_{\lambda \vdash (k,d)} \bigoplus_{\lambda' \vdash (k',d)} [V_\lambda \otimes V_{\lambda'}] \otimes (\lambda \otimes \lambda') = \bigoplus_{\mu} \bigoplus_{\lambda \vdash (k,d)} \bigoplus_{\lambda' \vdash (k',d)} m^{\mu}_{\lambda\lambda'} V_\mu \otimes (\lambda \otimes \lambda')$$

Now, consider the group restriction of the left side from $S_{k+k'}$ to the subgroup $S_k \times S_{k'} \subseteq S_{k+k'}$. Let $\mu \vdash (k+k')$ be a irreducible representation of $S_{k+k'}$. Under the group restriction

$$\text{Res}^{S_{k+k'}}_{S_k \times S_{k'}}[\mu] = \bigoplus_{\lambda \vdash k} \bigoplus_{\lambda' \vdash k'} B^{\lambda\lambda'}_{\mu} (\lambda \otimes \lambda')$$

where $B^{\lambda\lambda'}_{\mu}$ are the branching rules which count how many copies of the irreducible $\lambda \otimes \lambda'$ are contained in the restriction of μ. Branching rules for the symmetric group have been thoroughly studied [18]. Under group restriction from $S_{k+k'} \to S_k \times S_{k'}$, the isomorphism of vector spaces becomes an isomorphism of group representations. Under restriction

$$\bigoplus_{\mu \vdash (k+k',d)} V_\mu \otimes \mu \to \bigoplus_{\mu \vdash (k+k',d)} \bigoplus_{\tau \vdash k} \bigoplus_{\tau' \vdash k'} B^{\mu}_{\tau\tau'} V_\mu \otimes [\tau \otimes \tau']$$

Two representations are equivalent if and only if they have identical decomposition of irriducibles. This relation can only hold if, for any $\lambda \vdash k$ and $\lambda' \vdash k'$ the relation

$$V_\lambda \otimes V_{\lambda'} = \bigoplus_{\mu \vdash (k+k')} B^{\mu}_{\lambda\lambda'} V_\mu$$

holds. Thus, the tensor product of the λ and λ' irreducibles of $U(d)$ are completely determined by the branching rules of irreducible representations of the symmetric group. Branching rules of the symmetric group have been thoroughly studied in representation theory [18].

Specifically, Pieri's formula give a easy diagrammatic method for computing $B^{\mu}_{\lambda\lambda'}$.

IX. CHARACTERIZATION OF LIE ALGEBRA REPRESENTATIONS

The continuous structure of Lie groups allows for classification of their algebraic structure. In this section, we discuss the Killing form, Cartan sub-algebra and Weyl group/chamber. The Weyl group and Weyl chambers encode the symmetry of the root system, which is instrumental in understanding the geometry and representation theory of the group. The Weyl integration formula and the Harish-Chandra integral formula provide powerful tools for integrating functions over the group, linking harmonic analysis to representation theory. Using Dynkin diagrams, all compact Lie algebras can be classified into a finite list of types, giving a complete understanding of their structure and symmetries.

A. Killing Form

The Killing form is the first tool in the classification of Lie algebra representations. The Killing form K is a symmetric bi-linear form on a Lie algebra $\mathfrak{g}$. Specifically, K is defined as

$$K(X, Y) = \mathrm{Tr}[\mathrm{ad}(X)\mathrm{ad}(Y)]$$

Using the cyclic properties of the trace,

$$K(X, [Z, Y]) + K([Z, X], Y) = 0 \tag{5}$$

The Killing form is essentially unique. It is (up to multiplication) the only inner product satisfying the property 5. The Killing form can be written in terms of the structure constants f^k_{ij} as

$$K(A^i X_i, B^j X_j) = \sum_k f^k_{ij} f^k_{ji} A^i B^j$$

So that as an element of $\mathfrak{g}^\star \otimes \mathfrak{g}^\star$ the Killing form is given by

$$K = \sum_{km=1} f^k_{im} f^m_{jk} e^i \otimes e^j$$

where $\mathfrak{g}^\star = \mathrm{span}[e^i]^r_{i=1}$ is the dual space of $\mathfrak{g}$. Importantly, the Killing form is an Ad-invariant inner product,

$$\forall g \in G, \quad K(X, Y) = K(\mathrm{Ad}_g X, \mathrm{Ad}_g Y)$$

B. Cartan Sub-Algebra

A Cartan sub-algebra $\mathfrak{h}$ is a maximal commuting set of elements of $\mathfrak{g}$. A Cartan sub-algebra is closed under commutation and satisfies

$$\forall x, y \in \mathfrak{h}, \quad [x, y] = 0$$

The dimensions of $\dim \mathfrak{h} = r$ is called the rank of $\mathfrak{g}$. Let $\{h^i\}_{i=1}^r$ be a basis of $\mathfrak{h}$. The remaining elements of $\mathfrak{g}$ will be denoted as E^α where

$$\forall h \in \mathfrak{h}, \quad [h^i, E^\alpha] = \alpha^i E^\alpha$$

so that the E^α are eigenvectors of the h^i operators. The vectors $\alpha = (\alpha^1, \alpha^2, ..., \alpha^r)$ are called roots. The operator E^α is called the ladder operator associated to the root α. Let Φ denote all the roots of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ then decomposes as a direct sum of the Cartan sub-algebra and the roots

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} E^\alpha$$

Root systems have a reflection symmetry. Specifically, if α is a root, then $-\alpha$ is also a root as

$$[h^i, E^\alpha] = \alpha^i E^\alpha \implies [h^i, (E^\alpha)^\dagger] = -\alpha^i (E^\alpha)^\dagger$$

Using the Jacobi Identity, we have that

$$\forall h \in \mathfrak{h}, \quad [h^i, [E^\alpha, E^\beta]] = (\alpha + \beta)^i E^{\alpha+\beta}$$

thus, the commutator of two roots satisfies

$$[E^\alpha, E^\beta] = N_{\alpha,\beta} E^{\alpha+\beta} \text{ if } \alpha \neq -\beta$$
$$[E^\alpha, E^{-\alpha}] = \sum_{i=1}^r C_i(\alpha) h^i$$

where $N_{\alpha,\beta}$ and $C_i(\alpha)$ are constants. The constant $C_i(\alpha)$ can be determined using the Jacobi relation. We have that

$$[h^i, [E^\alpha, E^{-\alpha}]] + [E^\alpha, [E^{-\alpha}, h^i]] + [E^{-\alpha}, [h^i, E^\alpha]] = 0$$

Using the definition of roots, we have that

$$[h^i, [E^\alpha, E^{-\alpha}]] + 2\alpha^i [E^\alpha, E^{-\alpha}] = 0$$

Thus, $[E^\alpha, E^{-\alpha}]$ must be given by

$$[E^\alpha, E^{-\alpha}] = C(\alpha) \sum_{i=1}^r \alpha^i h^i$$

The root $\alpha(h) : \mathfrak{h} \to \mathbb{C}$ is the eigenvector of x in $[h, \cdot]$. Note that each root $\alpha : \mathfrak{h} \to \mathbb{C}$ can be viewed as an element of the dual space $\mathfrak{h}^\star$ of $\mathfrak{h}$. An orientation on a root system α is a choice of roots $\Phi^+ \subset \Phi$ such that either α or $-\alpha$ is contained in Φ^+, but not both. If the Lie algebra $\mathfrak{g}$ has an inner product, we can identify $\mathfrak{h}^\star$ with $\mathfrak{h}$. We can identify the dual $\mathfrak{h}^\star$ with $\mathfrak{h}$ via the canonical isomorphism $J : \mathfrak{h}^\star \to \mathfrak{h}$

$$J[x](y) = K(x, y)$$

where $K(\cdot, \cdot)$ is the Killing form on $\mathfrak{g}$. The Killing form induces a inner product on the root space. Let α and β be roots. We can then define the inner product on roots

$$(\alpha, \beta) = K(\sum_{i=1}^r \alpha^i h^i, \sum_{i=1}^r \beta^i h^i) = \sum_{i=1}^r \alpha^i \beta^i$$

The Killing form then defines a inner product in the dual space $\mathfrak{h}^\star$ via

$$(\alpha, \beta) = K(\alpha \cdot h, \beta \cdot h)$$

C. Weights

A weight vector $\lambda = (\lambda^1, \lambda^2, ..., \lambda^r)$ is a basis such that

$$\forall h^i, \quad h^i|\lambda\rangle = \lambda^i|\lambda\rangle$$

Using the commutation relations $[h^i, E^\alpha] = \alpha^i E^\alpha$, we have that

$$h^i[E^\alpha|\lambda\rangle] = (\lambda^i + \alpha^i)[E^\alpha|\lambda\rangle]$$

so that the operator E^α shifts the weight vector λ,

$$E^\alpha|\lambda\rangle \propto |\lambda + \alpha\rangle$$

The operator E^α is said to terminate the weight vector λ is there exists an integer $p \in \mathbb{Z}$ such that

$$(E^\alpha)^p|\lambda\rangle = 0$$

For finite representations, all the root operators E^α must terminate each weight vector $|\lambda\rangle$. Thus, we must have that

$$\frac{2(\alpha, \lambda)}{|\alpha|^2} \in \mathbb{Z}$$

This is called the Cartan relation. The Cartan relation forces the root and weight space to satisfy a set of natural geometric relations, allowing for a complete classification of simple Lie algebras.

D. Structures of Root Systems

The rank of the Cartan sub-algebra $\mathfrak{h}$ is in general much less than the dimension of the full Lie algebra $\mathfrak{g}$. Let $\{\beta_i\}_{i=1}^r$ be a basis of $\mathfrak{h}^\star$. Then, any root may be expanded as

$$\forall \alpha \in \Phi, \quad \alpha = \sum_{i=1}^{r} n_i \beta_i$$

where n_i are integers. Roots with the first non-zero $n_i > 0$ are called positive roots and denoted as Φ_+. A simple root is a root that cannot be written as the sum of two positive roots. The set of simple roots is denoted as Δ. There are exactly r simple roots. For any two simple roots, we define the Cartan matrix

$$\alpha_i, \alpha_j \in \Delta, \quad A_{ij} = \frac{2\langle \alpha_i, \alpha_j \rangle}{|\alpha_j|^2}$$

To each root $\alpha \in \Phi$, we associate a dual root $\alpha^\wedge$, defined as

$$\alpha^\wedge = \frac{2\alpha}{|\alpha|^2}$$

Using this definition, the Cartan matrix can be written as

$$A_{ij} = \langle \alpha_i, \alpha_j^\wedge \rangle$$

1. *Fundamental Weights*

The fundamental weights are defined as the normalized coroots with

$$(\omega_i, \alpha_j^\wedge) = \delta_{ij}$$

Any weight vector can be expanded in the fundamental weight basis as

$$\lambda = \sum_{i=1}^{r} \lambda_i \omega_i$$

where $\lambda_i = (\lambda, \alpha_i^\wedge)$ are called the Dynkin labels of λ. The Weyl vector ρ is defined as the sum of all fundamental weights

$$\rho = \sum_{i=1}^{r} \omega_i$$

E. Weyl Group

Consider the hyperplane defined by the equation

$$H_\alpha = \{ \ h \ | \ \langle \alpha, h \rangle > 0 \ \}$$

For any root $\alpha \in \Phi$, we can reflect around the hyperplane defined by H_α. The set of all reflections forms a group. Which is called the Weyl group W. Specifically, for any two roots β and α, the Weyl reflection of β with respect to α is given by

$$s_\alpha \beta = \beta - (\alpha^\wedge, \beta)\alpha$$

Because roots and weights live in the same space, the Weyl group also acts on weight vectors $|\lambda\rangle$ via

$$s_\alpha |\lambda\rangle = |\lambda\rangle - (\alpha^\wedge, \lambda)|\alpha\rangle$$

The Weyl group action on both weights and roots is unitary,

$$\begin{aligned} \text{Roots: } \forall w \in W, \ \forall \alpha, \alpha' \in \Phi \quad (\alpha, \alpha') &= (w\alpha, w\alpha') \\ \text{Weights: } \forall w \in W, \quad (\lambda, \lambda') &= (w\lambda, w\lambda') \end{aligned}$$

It will be useful to define the Fredenhall operator D_ρ as

$$D_\rho = \prod_{\alpha \in \Phi^+} (\exp(\alpha/2) - \exp(-\alpha/2))$$

using the definition of the Weyl group, this can be written in terms of the Weyl vector as

$$D_\rho = \sum_{w \in W} \eta(w) \exp(w\rho)$$

where $\eta(w) : W \to \pm 1$ is the sign function of W.

F. Weyl Chamber

The action of the Weyl group W on the root space splits the root space into $|W|$ isomorphic subspaces called chambers. The Weyl chamber defined as

$$W_c = \{ \ \lambda \ | \ \forall w \in W, \ \forall \alpha_i \in \Delta \quad (w\lambda, \alpha_i) \geq 0 \ \}$$

The discriminant function $\delta_{\mathfrak{g}}(x) : \mathfrak{h} \to \mathbb{C}$ is defined as

$$\forall x \in \mathfrak{h}, \quad \delta_{\mathfrak{g}}(x) = \prod_{\alpha \in \Phi^+} \langle \alpha, x \rangle$$

which is the products of the inner product of the Cartan element$x \in \mathfrak{h}$ with all positive roots.

G. Highest Weight Representations

A highest weight vector $|\lambda\rangle$ is a weight that is decimated by each positive root,

$$\forall \alpha \in \Phi^+, \quad E^\alpha |\lambda\rangle = 0$$

There is a bijection between highest weight representations and irreducible Lie algebra representations. Specifically, from a highest weight vector $|\lambda\rangle$, we can form the descendent states

$$\forall \alpha_i \in \Phi^+, \quad E^{-\alpha_1} E^{-\alpha_2} ... E^{-\alpha_m} |\lambda\rangle$$

Descent states form representations of the Lie algebra $\mathfrak{g}$. The set of all descendent states of the highest weight vector $|\lambda\rangle$ is denoted as L_λ.

The descendent states L_λ generate representation of the Lie algebra G. Specifically,

$$\begin{aligned} \text{Cartan Subgroup: } &\exp(\sum_{i=1}^{r}\theta_i h^i)|\lambda\rangle = \exp(\sum_{i=1}^{r}\theta_i\lambda^i)|\lambda'\rangle \\ \text{Lie Algebra: } &\exp(tE^\alpha)|\lambda'\rangle \in L_\lambda \end{aligned}$$

Thus, highest weight states generate representations of Lie groups. However, we have to keep track of both the multiplicities of the states in L_λ and be able to generate a basis for L_λ. Define the formal exponential $\exp(\mu)$ as a placeholder, where for all weights λ and λ',

$$\begin{aligned} \exp(\lambda+\lambda') &= \exp(\lambda)\exp(\lambda') \\ \exp(\lambda)(\lambda') &= \exp((\lambda,\lambda')) \end{aligned}$$

The character of the highest weight representation $|\lambda\rangle$ is then defined as

$$\chi_\lambda = \sum_{\lambda'\in L_\lambda}\text{Mult}_\lambda[\lambda']\exp(\lambda')$$

where the integer $\text{Mult}_\lambda[\lambda']$ counties the number of copies of the descendent state $|\lambda'\rangle$ in the $|\lambda\rangle$ highest weight representation. In general, calculating the Lie algebra characters is difficult. However, it can be show that the Freudenthal operator satisfies

$$D_\rho\chi_\lambda = D_{\rho+\lambda}$$

Thus, we have that

$$\chi_\lambda = \frac{D_{\rho+\lambda}}{D_\rho} \tag{6}$$

This 6 is called the Weyl character formula. Using 6, the dimension of a highest weight representation $|\lambda\rangle$ is given by

$$d_\lambda = \dim\lambda = \prod_{\alpha\in\Phi^+}\frac{(\rho+\lambda,\alpha)}{(\rho,\alpha)}$$

X. HARISH-CHANDRA INTEGRAL FORMULA

The Harish-Chandra integrals were discovered by Harish-Chandra in his development of the theory of harmonic analysis on semi-simple Lie groups. The HCIZ integrals [17] are a special case of the more general Harish-Chandra formula. Let G be a semi-simple group. Let $\text{Ad}: G\to\text{Aut}(\mathfrak{g})$ be the adjoint operator on G. Let W be the Weyl group of G. Let $\langle\cdot,\cdot\rangle : \mathfrak{g}\times\mathfrak{g}\to\mathbb{C}$ be any Ad-invariant inner product on $\mathfrak{g}$. Then, the Harish-Chandra formula evaluates integrals of the form

$$\int_{g\in G} dg\ \exp(\langle\text{Ad}_g(x),y\rangle)$$

in terms of summations over the the Weyl group W. Specifically,

$$\int_{g\in G} dg\ \exp(\langle\text{Ad}_g(x),y\rangle) = \frac{1}{\text{Vol}(W)}\sum_{w\in W}\text{sign}(w)\exp(\langle w(x),y\rangle)$$

where $w(x)$ is the lattice vector of x on W and $\text{sign}: W\to\pm 1$ is the sign function.

XI. BOCHNER THEOREMS

Bochner's theorem is a celebrated result in harmonic analysis [3, 12]. Modifications of Bochner's theorem have been used in high dimensional statistics and dimensionality reduction [1, 24]. The Bochner theorem on abelian groups was used in [24] to construct the Fourier random features. We review Bochner theorem on abelian groups, as stated in [24].

A. Bochner's Theorem on Abelian Groups

Let G be an Abelian group. All irreducible representations of G are one dimensional. The set of irreducible representations of G, denoted as $\hat{G}$, is again a group. The set of irreducible representations of the dual group $\hat{G}$ is isomorphic to G, so that $\hat{\hat{G}} = G$ This observation forms the basis of Pontryagin duality [23]. Bochner's theorem on abelian groups states the following: Let G be locally compact abelian group. Let $f : G \to \mathbb{R}$ be a normalized function on G, $\int_{g\in G} dg\ f(g) = 1$ which is positive definite, so that

$$\forall g_1, g_2, ..., g_k \in G,\ \ \forall v_1, v_2, ..., v_k \in \mathbb{C} \quad \sum_{ij=1}^{k} f(g_i^{-1} g_j)\bar{v}_i v_j \geq 0$$

Then, there exists a probability measure $d\mu_f$ on the dual group $\hat{G}$ such that

$$f(g) = \int_{\omega\in\hat{G}} d\mu_f(\omega)\ \omega(g)$$

Thus, normalized functions on G have positive definite measure on the the Pontryagin dual $\hat{G}$. Any positive definite normalized left-invariant function $f(g_1, g_2) = f(hg_1, hg_2)$ can then be written as

$$f(g_1, g_2) = f(g_1^{-1} g_2, e) = \int_{\omega\in\hat{G}} d\mu_f(\omega)\ \omega(g_1^{-1} g_2) = \int_{\omega\in\hat{G}} d\mu_f(\omega)\ \omega(g_1^{-1})\omega(g_2)$$

Then, using the positively of the measure $d\mu_f(\omega)$, we can write this as

$$f(g_1, g_2) = f(g_2^{-1} g_1, e) = \mathbb{E}_{\omega\sim\mu_f}[\omega(g_2^{-1})\omega(g_1)]$$

which is the random features expansion for $f(\cdot, \cdot)$. Because of this, random features for kernels can be constructed by sampling from the probability measure μ_f on $\hat{G}$. This observation was used in [24] to construct the random Fourier features, and is the basis of all random features methods (cite). However, this version XI A of Bochner's theorem only holds for functions f defined on a abelian group G. What if we have a function f that is defined on a homogeneous space $X = G/H$ of a non-commutative group G? This is not an academic question: The softmax kernel in attention is defined on $\mathbb{R}^d$ which is a homogeneous space of the non-commutative group $E(d) = O(d) \rtimes \mathbb{R}^d$. Can we develop a analogy for Bochner's theorem on homogeneous spaces of non-commutative groups?

B. Bochner's Theorem on Compact Groups

Bochner proved an additional related theorem for compact groups [2]. Specifically, let G be a compact group, not necessarily non-commutative. Let $f : G \times G \to \mathbb{C}$ be a left G-invariant $f(g_1, g_2) = f(hg_1, hg_2)$, positive definite function on G that satisfies following property:

$$\forall v_1, v_2, ..., v_k \in \mathbb{C},\ \ \forall g_1, g_2, ..., g_n \in G, \quad \sum_{ij=1}^{k} v_i \bar{v}_j f(g_i, g_j) \geq 0 \tag{7}$$

Consider the Fourier expansion of f,

$$f(g_1, g_2) = f(g_2^{-1} g_1, e) = \sum_{g\in\hat{G}} d_\rho \mathrm{Tr}[\hat{f}^\rho \rho(g_1^{-1} g_2)]$$

Then, Bochner's theorem [2] states that each of the matrix expansion coefficients $\hat{f}^\rho$ are positive definite, $\hat{f}^\rho \succcurlyeq 0$.

1. *Non-Commutative Random Features*

We can use Bochner's theorem [2] on compact groups to construct random features approximations. Specifically, Each of the $\hat{f}^\rho$ matrices can be diagonlized as $\hat{f}^\rho = U^\rho \Lambda^\rho (U^\rho)^\dagger$ where Λ^ρ is a diagonal matrix with positive semi-definite entries. Bochners theorem on compact groups allows us to write positive definite functions on compact groups as expectations of random features. Specifically, suppose that $f : G \times G \to \mathbb{R}$ is a normalized left G-invariant positive definite function satisfying 7. Then, using Bochner's theorem on compact groups, we may write

$$f(g_1, g_2) = f(g_2^{-1}g_1, e) = \sum_{\rho\in\hat{G}} d_\rho \mathrm{Tr}[U^\rho \Lambda^\rho (U^\rho)^\dagger \rho(g_1^{-1}g_2)]$$

Thus, we may write

$$f(g_2^{-1}g_1, e) = \sum_{\rho\in\hat{G}} d_\rho \mathrm{Tr}[\rho(g_1) U^\rho \Lambda^\rho (U^\rho)^\dagger \rho(g_2^{-1})]$$

If we define a set of d_ρ vectors, each of dimension d_ρ with

$$\forall ij \in \{1, 2, ..., d_\rho\}, \quad \Phi_i^\rho(g)_j = \sqrt{d_\rho \hat{K}^\rho} \sum_{k=1}^{d_\rho} \bar{U}_{ik}^\rho \rho(g)_{kj}$$

Then, we have that

$$f(g_1, g_2) = \sum_{\rho\in\hat{G}} \sum_{i=1}^{d_\rho} \Phi_i^\rho(g)^T \Phi_i^\rho(g)$$

Thus, the positivity of $\hat{f}^\rho$ guaranteed by Bochner's theorem allows for a random features expansion of any positive definite function on compact groups. On compact groups we don't even need random features. Specifically, we can define the deterministic vector

$$\Phi(g) = \mathrm{Concat}_{\rho\in\hat{G}}[\mathrm{Concat}_{1\le m\le d_\rho}[\sqrt{d_\rho K^\rho} U_m^\rho \rho(g)]]$$

then,

$$\forall g_1, g_2 \in G, \quad K(g_1, g_2) = \Phi(g_1)^T \Phi(g_2)$$

holds exactly. For compact Lie groups, this sum will be infinite and must be truncated at some maximum harmonic ℓ,

$$\Phi_\ell(g) = \mathrm{Concat}_{\rho\in\hat{G}}^{d_\rho\le\ell}[\mathrm{Concat}_{1\le m\le d_\rho}[\sqrt{d_\rho K^\rho} U_m^\rho \rho(g)]]$$

The error in this truncation is highly controlled. Specifically, using the Parsifal-Plancheral theorem **??**,

$$||K - \Phi_\ell^T \Phi_\ell||_{L^2[G]} = \sum_{\rho\in\hat{G}, d_\rho>\ell} d_\rho^2 ||\hat{K}^\rho||_F^2$$

so the validity of the approximation is determined by how quickly the Fourier coefficients $\hat{K}^\ell$ decay to zero.

2. *Example: Non-Commutative Random Features on G*

Let $\sigma \in \hat{G}$ be an irreducible representation. Consider the F-norm kernel of the σ representation,

$$K_\sigma(g_1, g_2) = ||\sigma(g_1) - \sigma(g_2)||_F^2$$

The Fourier coefficients are given by

$$\hat{K}_\sigma^\rho = \delta_\sigma^\rho \hat{K}^\sigma \mathbb{1}_{d_\sigma}$$

where $\hat{K}^\rho \in \mathbb{R}^+$. For this choice of kernel function, the Fourier matrix $\hat{K}_\sigma^\rho$ is proportional to the identity when $\rho = \sigma$ and zero otherwise. The random features decomposition is then

$$\forall g \in G, \quad \Phi_\omega^\rho(g)_i = \sqrt{d_\rho} \delta_\sigma^\rho \sigma(g)_{i\omega}$$

where the index $\omega \in \{1, 2, ..., d_\sigma\}$ is chosen uniformly at random. Note that, independent of the random variable ω,

$$\forall h, g \in G, \quad \Phi_\omega^\rho(h \cdot g)_i = \sum_{j=1}^{d_\rho} \rho(h)_{ij} \Phi_\omega^\rho(g)_j$$

so that independent of the random variable ω, $\Phi_\omega^\rho(g)_i$ transforms in the ρ representation of G.

C. Bochner's Theorem on Homogeneous Spaces of Compact Groups

- Expand on proof of this, Still too terse,

The results of XI B can be easily generalized to homogeneous spaces of compact groups. Let G be a compact group and let $H \subseteq G$ be a subgroup of G. Let $X = G/H$ be a homogeneous space of G. Suppose that f is a positive definite, left G-invariant, $f(gx, gy) = f(x, y)$ normalized function on $X \times X$ so that $\int_{x,y \in X} dxdy\ f(x, y) = 1$ and

$$\forall v_1, v_2, ..., v_k \in \mathbb{C}, \ \ \forall x_1, x_2, ..., x_k \in X, \quad \sum_{j=1}^{k} v_i \bar{v}_j f(x_i, x_j) \geq 0$$

Then, using a result of [19], we may expand f as

$$f(x, y) = \sum_{\rho \in \hat{G}} \mathrm{Tr}[\hat{f}^{\rho} \rho([g(y)^{-1} g(x)]^{-1})]$$

where the coefficients $\hat{f}^{\rho}$ have additional sparsity constraints (see Proposition 1 in [19]) and $g(x)$ is the G representative of $x \in X = G/H$. Again, applying Bochner's Theorem on Compact Groups XI B, each of the matrices $\hat{f}^{\rho} \succcurlyeq 0$ is positive semi-definite.

XII. BEYOND COMPACT GROUPS: REPRESENTATION THEORY ON LOCALLY COMPACT GROUPS

Compact groups always have irreducible representations of finite dimension. Locally compact groups do not have to satisfy this property. Specifically, locally compact groups may have irreducible representations that live in a Hilbert space of infinite dimensions. Hilbert spaces [10] are generalizations of finite dimensional vector spaces to vector spaces that may be uncountably infinite-dimensional. Formally, a Hilbert space is a vector space V equipped with an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ such that the $(\langle \cdot, \cdot \rangle, V)$ is a complete metric space [10]. Examples of Hilbert spaces are square integrable function spaces with inner product,

$$\langle f, g \rangle = \int dx\ \bar{f}(x) g(x)$$

Locally compact groups have unitary irreducible representations $\hat{G}$ that carry both point and continuous indices. Fortunately, locally compact groups behave very similar to compact groups. Let $\mathcal{H}$ be a Hilbert space. A unitary Lie representation of $(\rho, \mathcal{H})$ is a representation such that

$$\forall h, k \in \mathcal{H}, \quad g \to \langle h | \rho(g) | k \rangle$$

is a continuous function of g. Locally compact groups, although not compact, satisfy the following key property: the Gel'fand–Raikov theorem [29] then states that points of G are *separated* by irreducible unitary representations of G, i.e. for any two group elements $g, g' \in G$, there exist a unitary Lie representation $(\rho, \mathcal{H})$ such that

$$\rho(g) \neq \rho(g')$$

Thus, the matrix elements $\langle h | \rho(g) | k \rangle$ are dense in the space of square integrable functions. In other words, for any compact subset of the group, if $f \in \mathbb{C}$ then there is an expansion of $f(g)$ in terms of the matrix elements $\langle h | \rho(g) | k \rangle$. The actual statement of Gel'fand-Raikov [29] is "For every locally compact group, there exists a complete system of irreducible unitary representations." It should also be noted that the Parseval-Plancheral theorem can also be generalized to locally compact groups [25]. Using the Gel'fand-Raikov theorem [29], the set of matrix elements of unitary irreducible representations is dense in the space of square-integrable functions on a locally compact group. Harmonic functions are defined as the overlap of the irreducible representation ℓ matrix elements $|\ell k\rangle$ in the position basis as,

$$\forall x \in X, \quad Y_{\ell}(x)_k = \langle x | \ell k \rangle$$

where $|x\rangle$ is the position ket. In general, harmonics have an index ρ that is discrete and an index k that is continuous. The harmonics are orthogonal in both the point and continuous index,

$$\int_{x\in X} dx\ Y_{k\rho}(x)Y_{k'\tau}(x) = \delta_{\rho\sigma}\delta^{(d)}(k-k')$$

where $\delta^{(d)}(x)$ is the delta function in d-dimensions. Furthermore, the harmonics form a complete basis of $L^2[X]$, so that

$$\forall f \in L^2[X], \quad f(x) = \sum_{\rho}\int dk\ \hat{f}_\rho(k)Y_{\rho,k}(x)$$

The Fourier coefficients are given by

$$\hat{f}_\rho(k) = \int_{x\in X} dx\ f(x)Y_{\rho,k}(x)$$

where the Fourier coefficients $\hat{f}_\rho(k)$ carry discrete index ρ and continuous index k.

Acknowledgments

Owen Howell thanks his cat Boswell for useful discussions. Owen Howell thanks the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) for financial support.

[1] Raj Agrawal, Trevor Campbell, Jonathan H. Huggins, and Tamara Broderick. Data-dependent compression of random features for large-scale kernel approximation, 2019.

[2] S. Bochner. Hilbert distances and positive definite functions. *Annals of Mathematics*, 42(3):647–656, 1941.

[3] Saloman Bochner. *Harmonic Analysis and the Theory of Probability.* University of California Press, Berkeley, 1955.

[4] Michael Brodskiy and Owen L. Howell. k-fold gaussian random matrix ensembles i: Forcing structure into random matrices, 2024.

[5] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, 2021.

[6] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Harmonic Analysis on Finite Groups: Representation Theory, Gelfand Pairs and Markov Chains.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[7] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Harmonic Analysis on Finite Groups: Representation Theory, Gelfand Pairs and Markov Chains.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[8] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Induced representations and Mackey theory*, page 399–425. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2018.

[9] Krzysztof Choromanski, Valerii Likhosherstov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Davis, Afroz Mohiuddin, Lukasz Kaiser, David Belanger, Lucy Colwell, and Adrian Weller. Rethinking attention with performers, 2022.

[10] L. Debnath and P. Mikusinski. *Introduction to Hilbert Spaces with Applications.* Elsevier Science, 2005.

[11] Philippe Di Francesco, Pierre Mathieu, and David Sénéchal. *Conformal field theory.* Graduate texts in contemporary physics. Springer, New York, NY, 1997.

[12] G.B. Folland. *A Course in Abstract Harmonic Analysis.* Studies in Advanced Mathematics. Taylor & Francis, 1994.

[13] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Graduate Texts in Mathematics. Springer International Publishing, 2015.

[14] Owen Howell, Haoen Huang, and David Rosen. Multi-irreducible spectral synchronization for robust rotation averaging, 2023.

[15] Owen Howell, David Klee, Ondrej Biza, Linfeng Zhao, and Robin Walters. Equivariant single view pose prediction via induced and restricted representations, 2023.

[16] Haojie Huang, Owen Howell, Dian Wang, Xupeng Zhu, Robin Walters, and Robert Platt. Fourier transporter: Bi-equivariant robotic manipulation in 3d, 2024.

[17] C. Itzykson and J. B. Zuber. The Planar Approximation. 2. *J. Math. Phys.*, 21:411, 1980.

[18] G.D. James and A. Kerber. *The Representation Theory of the Symmetric Group.* Encyclopedia of mathematics and its applications. Addison-Wesley Publishing Company, Advanced Book Program, 1981.

[19] Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups, 2018.

[20] Antoine Neven, Jose Carrasco, Vittorio Vitale, Christian Kokail, Andreas Elben, Marcello Dalmonte, Pasquale Calabrese, Peter Zoller, Benoit Vermersch, Richard Kueng, and Barbara Kraus. Symmetry-resolved entanglement detection using partial transpose moments. *npj Quantum Information*, 7(1), oct 2021.

[21] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[22] Silvia Pappalardi, Laura Foini, and Jorge Kurchan. Eigenstate thermalization hypothesis and free probability. *Physical Review Letters*, 129(17), October 2022.

[23] L. Pontrjagin. The theory of topological commutative groups. *Annals of Mathematics*, 35(2):361–388, 1934.

[24] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Neural Information Processing Systems*, 2007.

[25] I. E. Segal. An extension of plancherel's formula to separable unimodular groups. *Annals of Mathematics*, 52(2):272–292, 1950.

[26] J. P. Serre. Groupes finis, 2005.

[27] HERMANN WEYL. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, 1966.

[28] HERMANN WEYL. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, 1966.

[29] Hisaaki Yoshizawa. Unitary representations of locally compact groups. Reproduction of Gelfand-Raikov's theorem. *Osaka Mathematical Journal*, 1(1):81 – 89, 1949.

[30] Xiao-Dong Yu, Satoya Imai, and Otfried Gühne. Optimal entanglement certification from moments of the partial transpose. *Physical Review Letters*, 127(6), aug 2021.

[31] A. Zee. *Group Theory in a Nutshell for Physicists*. In a Nutshell. Princeton University Press, 2016.

[32] A. Zee. *Group Theory in a Nutshell for Physicists*. In a Nutshell. Princeton University Press, 2016.

[33] Linfeng Zhao, Owen Howell, Xupeng Zhu, Jung Yeon Park, Zhewen Zhang, Robin Walters, and Lawson L. S. Wong. Equivariant action sampling for reinforcement learning and planning, 2024.

Appendix A: Representation Theory of Unitary Group $U(d)$

The representation theory of the group $U(d)$ was worked out in the early 1900s by Jacobi, Schur and Weyl, among others. The representation theory of the group $U(d)$ is especially elegant and is intimately related to the representation theory of the symmetric group. The unitary group $U(d)$ is both semi-simple and compact so the set of irreducible representations of $U(d)$ are countably infinite. Let $\lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ be an integer partition with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m$. Characters of irreducible representations are given by

$$s_\lambda(z_1, z_2, ..., z_m) = \chi_\lambda(z) : (\mathbb{C}^\times)^m \to \mathbb{C}$$

where the s_λ are called called Schur functions. Define the function

$$a_{\lambda_1,\lambda_2,...,\lambda_m}(z_1, z_2, ..., z_m) = \det \begin{bmatrix} z_1^{\lambda_1+m-1} & z_2^{\lambda_1+m-1} & ... & z_n^{\lambda_1+m-1} \\ z_1^{\lambda_2+m-2} & z_2^{\lambda_2+m-2} & ... & z_n^{\lambda_2+m-2} \\ ... & ... & ... & ... \\ z_1^{\lambda_n} & z_2^{\lambda_n} & ... & z_n^{\lambda_n} \end{bmatrix}$$

The Schur function is the defined by

$$s_\lambda(z_1, z_2, ..., z_m) = \frac{a_\lambda(z_1, z_2, ..., z_m)}{\Delta(z_1, z_2, ..., z_m)}$$

where $\Delta(z)$ is the Vandermode determinant.

Appendix B: Multi-Linear Algebra

We briefly review some multi-linear algebra concepts and operations on tensor product spaces. We specifically discuss partial transpose and partial conjugation, which are some standard tools in quantum information theory [21].

1. Partial Trace

The partial trace is a standard tool in quantum information theory [21]. Let $H = H_A \otimes H_B$ be a Hilbert space composed of the H_A and H_B Hilbert spaces. Let O be an operator defined on W. The partial trace of an operator on the H_A or H_B subspace is then defined as

$$O^{(A)} = \mathrm{Tr}_B[O], \quad O^{(B)} = \mathrm{Tr}_A[O]$$

respectively, where the matrix elements of the partially traced operators are defined as

$$O^{(A)}_{ij} = \sum_{k=1}^{d_B} O_{ik,jk}, \quad O^{(B)}_{ij} = \sum_{k=1}^{d_A} O_{ki,kj}$$

An operator O is said to be separable if $O = O_A \otimes O_B$ factorizes. Partial traces of separable operators satisfy

$$O^{(A)} = \mathrm{Tr}_B[O] = \mathrm{Tr}[O_B]O_A, \quad O^{(B)} = \mathrm{Tr}_A[O] = \mathrm{Tr}[O_A]O_B$$

A generic operator is not separable. However, via the operator-Schmidt decomposition.

Theorem 2 (Operator Schmidt-Decomposition)**.** *Let O be an operator defined on the $V \otimes V$ tensor product space. The operator O can always be written as*

$$O = \sum_{\ell=1}^{N_O} p_\ell A_\ell \otimes B_\ell$$

where p_ℓ are positive real numbers and the operators A_ℓ and B_ℓ are orthogonal on the V subspaces,

$$Tr[A_\ell^\dagger A_{\ell'}] = \delta_{\ell\ell'} = Tr[B_\ell^\dagger B_{\ell'}]$$

the integer N_O (the rank of the matrix) counts the minimum number of tensor product operators needed to decompose O. N_O is called the Schmidt number of the operator O.

The partial trace operation satisfies a uniqueness property.

Theorem 3 (Uniqueness of Partial Trace)**.** *The partial trace is the unique linear map*

$$Tr_B : L(A \otimes B) \to L(A) \tag{B1}$$

that satisfies the property

$$\forall H_B \in L(B),\ \ \forall H_A \in L(A), \quad Tr_B[H_A \otimes H_B] = Tr[H_B]H_A$$

2. Partial Transpose and Partial Conjugation

Let V be a vector space over $\mathbb{C}$ of dimension d. Let $V \otimes V$ be the vector space which is a tensor product of V with itself. The partial transpose and partial conjugate are often used tools in quantum information theory [20, 30].

a. Partial Transpose

Let $|ij\rangle$ be a set of basis elements of $V \otimes V$. We will denote P_1 as the partial transpose on the first copy of the V subspace and P_2 as the partial transpose on the second copy of the V subspace. Using Bra-Ket notation, We have that,

$$\langle ij|X^{P_1}|k\ell\rangle = \langle kj|X|i\ell\rangle \quad \langle ij|X^{P_2}|k\ell\rangle = \langle i\ell|X|kj\rangle$$

The action of P_1 and P_2 is commutative. The action of P_1 followed by P_2 (or vise versa) returns the matrix transpose. For any matrix X on $V \otimes V$,

$$(X^{P_1})^{P_2} = X^T = (X^{P_2})^{P_1}$$

A separable operator is one that can be written as the tensor product of two operators. Let X be a separable operator with $X = X_1 \otimes X_2$. Then the partial transpose of the operator X on the i-th tensor product subspace P_i is defined as

$$X^{P_1} = X_1^T \otimes X_2 \quad X^{P_2} = X_1 \otimes X_2^T$$

Any matrix X on $V \otimes V$ can always be written as a sum of separable operators (Need a Cite here)

$$X = \sum X_i^{(1)} \otimes X_i^{(2)}$$

and the partial transpose of X given by

$$X^{P_1} = \sum (X_i^{(1)})^T \otimes X_i^{(2)} \quad X^{P_2} = \sum X_i^{(1)} \otimes (X_i^{(2)})^T$$

The transpose of a matrix satisfies the property,

$$(XY)^T = Y^T X^T$$

A similar rule holds for the partial transpose. We have that

$$(XY)^{P_1} = (Y^T X^T)^{P_2} \quad (XY)^{P_2} = (Y^T X^T)^{P_1}$$

If $X = X_1 \otimes X_2$ is a separable operator, this identity can be further simplified. We have that,

$$(XY)^{P_1} = (\mathbb{1}_d \otimes X_2) Y^{P_1} (X_1^T \otimes \mathbb{1}_d) \quad (YX)^{P_1} = (X_1^T \otimes \mathbb{1}_d) Y^{P_1} (\mathbb{1}_d \otimes X_2)$$

b. Partial Conjugation

Partial Conjugation is the complex version of Partial Transposition. Let X be a separable operator with $X = X_1 \otimes X_2$. Then the partial conjugate of the operator X on the i-th tensor product subspace C_i is defined as

$$X^{C_1} = X_1^\dagger \otimes X_2 \quad X^{C_2} = X_1 \otimes X_2^\dagger$$

Any matrix X on $V \otimes V$ can always be written as a sum of separable operators

$$X = \sum X_i^{(1)} \otimes X_i^{(2)}$$

and the partial conjugate of X given by

$$X^{C_1} = \sum (X_i^{(1)})^\dagger \otimes X_i^{(2)} \quad X^{C_2} = \sum X_i^{(1)} \otimes (X_i^{(2)})^\dagger$$

The action of C_1 and C_2 is commutative. The action of C_1 followed by C_2 (or vise versa) returns the matrix conjugate. For any matrix X on $V \otimes V$,

$$(X^{C_1})^{C_2} = X^\dagger = (X^{C_2})^{C_1}$$

The conjugate of a matrix satisfies the property,

$$(XY)^\dagger = Y^\dagger X^\dagger$$

A similar rule holds for the partial transpose. We have that

$$(XY)^{C_1} = (Y^\dagger X^\dagger)^{C_2} \quad (XY)^{C_2} = (Y^\dagger X^\dagger)^{C_1}$$

If $X = X_1 \otimes X_2$ is a separable operator, this identity can be further simplified. We have that,

$$(XY)^{C_1} = (\mathbb{1}_d \otimes X_2) Y^{C_1} (X_1^\dagger \otimes \mathbb{1}_d) \quad (YX)^{C_1} = (X_1^\dagger \otimes \mathbb{1}_d) Y^{C_1} (\mathbb{1}_d \otimes X_2)$$

---

[1] Raj Agrawal, Trevor Campbell, Jonathan H. Huggins, and Tamara Broderick. Data-dependent compression of random features for large-scale kernel approximation, 2019.

[2] S. Bochner. Hilbert distances and positive definite functions. *Annals of Mathematics*, 42(3):647–656, 1941.

[3] Saloman Bochner. *Harmonic Analysis and the Theory of Probability.* University of California Press, Berkeley, 1955.

[4] Michael Brodskiy and Owen L. Howell. k-fold gaussian random matrix ensembles i: Forcing structure into random matrices, 2024.

[5] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, 2021.

[6] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Harmonic Analysis on Finite Groups: Representation Theory, Gelfand Pairs and Markov Chains.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[7] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Harmonic Analysis on Finite Groups: Representation Theory, Gelfand Pairs and Markov Chains.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[8] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Induced representations and Mackey theory*, page 399–425. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2018.

[9] Krzysztof Choromanski, Valerii Likhosherstov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Davis, Afroz Mohiuddin, Lukasz Kaiser, David Belanger, Lucy Colwell, and Adrian Weller. Rethinking attention with performers, 2022.

[10] L. Debnath and P. Mikusinski. *Introduction to Hilbert Spaces with Applications.* Elsevier Science, 2005.

[11] Philippe Di Francesco, Pierre Mathieu, and David Sénéchal. *Conformal field theory.* Graduate texts in contemporary physics. Springer, New York, NY, 1997.

[12] G.B. Folland. *A Course in Abstract Harmonic Analysis.* Studies in Advanced Mathematics. Taylor & Francis, 1994.

[13] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Graduate Texts in Mathematics. Springer International Publishing, 2015.

[14] Owen Howell, Haoen Huang, and David Rosen. Multi-irreducible spectral synchronization for robust rotation averaging, 2023.
[15] Owen Howell, David Klee, Ondrej Biza, Linfeng Zhao, and Robin Walters. Equivariant single view pose prediction via induced and restricted representations, 2023.
[16] Haojie Huang, Owen Howell, Dian Wang, Xupeng Zhu, Robin Walters, and Robert Platt. Fourier transporter: Bi-equivariant robotic manipulation in 3d, 2024.
[17] C. Itzykson and J. B. Zuber. The Planar Approximation. 2. *J. Math. Phys.*, 21:411, 1980.
[18] G.D. James and A. Kerber. *The Representation Theory of the Symmetric Group*. Encyclopedia of mathematics and its applications. Addison-Wesley Publishing Company, Advanced Book Program, 1981.
[19] Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups, 2018.
[20] Antoine Neven, Jose Carrasco, Vittorio Vitale, Christian Kokail, Andreas Elben, Marcello Dalmonte, Pasquale Calabrese, Peter Zoller, Benoit Vermersch, Richard Kueng, and Barbara Kraus. Symmetry-resolved entanglement detection using partial transpose moments. *npj Quantum Information*, 7(1), oct 2021.
[21] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
[22] Silvia Pappalardi, Laura Foini, and Jorge Kurchan. Eigenstate thermalization hypothesis and free probability. *Physical Review Letters*, 129(17), October 2022.
[23] L. Pontrjagin. The theory of topological commutative groups. *Annals of Mathematics*, 35(2):361–388, 1934.
[24] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Neural Information Processing Systems*, 2007.
[25] I. E. Segal. An extension of plancherel's formula to separable unimodular groups. *Annals of Mathematics*, 52(2):272–292, 1950.
[26] J. P. Serre. Groupes finis, 2005.
[27] HERMANN WEYL. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, 1966.
[28] HERMANN WEYL. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, 1966.
[29] Hisaaki Yoshizawa. Unitary representations of locally compact groups. Reproduction of Gelfand-Raikov's theorem. *Osaka Mathematical Journal*, 1(1):81 – 89, 1949.
[30] Xiao-Dong Yu, Satoya Imai, and Otfried Gühne. Optimal entanglement certification from moments of the partial transpose. *Physical Review Letters*, 127(6), aug 2021.
[31] A. Zee. *Group Theory in a Nutshell for Physicists*. In a Nutshell. Princeton University Press, 2016.
[32] A. Zee. *Group Theory in a Nutshell for Physicists*. In a Nutshell. Princeton University Press, 2016.
[33] Linfeng Zhao, Owen Howell, Xupeng Zhu, Jung Yeon Park, Zhewen Zhang, Robin Walters, and Lawson L. S. Wong. Equivariant action sampling for reinforcement learning and planning, 2024.