The Peter-Weyl Theorem:
Harmonic analysis and representation theory of compact groups

Ben McDonough
Math 480

February 19 and 21, 2024

Contents

Foreword

The Peter-Weyl theorem is a collection of fundamental results in harmonic analysis for compact topological groups which are not necessarily Abelian. This theorem was originally proven by Hermann Weyl with his student Fritz Peter in 1927. The Peter-Weyl theorem leverages the existence of the left-invariant Haar measure [4] to establish a remarkable generalization of results about the left-regular representations of finite groups to a compact topological group G. The main statement of the Peter-Weyl theorem is the following:

The matrix coefficients of continuous, finite-dimensional irreducible representations of G are dense in $L^2(G)$.[1]

This is the basic content of the theorem. The matrix coefficients themselves form a ring $C_{\text{alg}}(G)$ which satisfy many incredible algebraic *and* analytic properties, so we will spend some time studying this ring. In particular, if $C(G)^{\text{fin}}$ is the subspace of continuous functions on G that generate a finite-dimensional subspace under the left-action of G, it turns out that

$$C_{\text{alg}}(G) = C(G)^{\text{fin}}$$

This remarkable relationship is the core of the connection between the density of $C_{\text{alg}}(G)$ in $L^2(G)$ and the representation theory of G. Our proof strategy using the Stone-Weierstrass Theorem 13 highlights the algebraic structure of $C_{\text{alg}}(G)$.

We will also prove several statements which are equivalent to the Peter-Weyl theorem in different settings:

(1) The space $L^2(G)$ under the left-action of G decomposes as $L^2(G) \cong \widehat{\bigoplus}_{V \in \widehat{G}} V^{\oplus \dim V}$, where $\widehat{G}$ is the set of finite-dimensional non-isomorphic representations of G.[2]

(2) If H is a Hilbert space representation of G, then H^{fin}, the subspace of H consisting of vectors that generate a finite-dimensional subspace under the action of G, is dense in H.

(3) Any unitary representation of G over a Hilbert space H decomposes as $H = \widehat{\bigoplus}_{V \in \widehat{G}} V^{\oplus n_V}$, where n_V is the multiplicity of V in H.

(4) If G is a compact Lie group, then it is isomorphic to a subset of $\mathrm{GL}(n, \mathbb{C})$ for some finite n.

(5) Under the bi-action of $G \times G$ by left- and right-translation, we have $L^2(G) \cong \widehat{\bigoplus}_{V \in \widehat{G}} V \otimes V$.

These are all sometimes referred to as the Peter-Weyl theorem. The first and third statements say that an analogue of complete reducibility holds for infinite-dimensional representations of all compact topological groups. The first three statements are equivalent to the Peter-Weyl theorem in its basic form. The fourth statement only applies to Lie groups, but in this setting it also equivalent to the Peter-Weyl theorem. The last statement is almost just a restatement of the first, but it turns out to generalize Fourier analysis to non-Abelian groups. We will develop this generalization by proving the Parseval-Plancherel Formula, the Fourier Inversion Formula, and the Convolution Theorem.

Preliminaries

With all this said, let us begin by recalling some foundational definitions and results of analysis and representation theory.

Hilbert space preliminaries

For a finite group G, we have only finite-dimensional irreducible representationions. This may not be the case for an infinite group. In particular, the Peter-Weyl theorem only concerns infinite-dimensional representations in a **Hilbert space.**

Definition 1 (Hilbert Space)**.** A **Hilbert space** is a vector space equipped with a bilinear form $\langle -, - \rangle : H \times H \to \mathbb{C}$ which is linear in the first argument, skew-symmetric, and positive definite, and H is complete with respect to the norm induced by this form.

[1]The definition of matrix coefficients is Def. 21

[2]The symbol $\widehat{\bigoplus}$ refers to the completion of the direct sum, which is explained precisely in 8.

Proposition 1 (Properties of inner product and norm)**.** If H is a Hilbert space equipped with an inner product $\langle -, - \rangle$, then the following three identities are satisfied:

1. (Cauchy-Schwarz) $|\langle v, w \rangle| \leq \|v\| \|w\|$
2. (Triangle inequality) $\|v + w\| \leq \|v\| + \|w\|$
3. (Parallelogram) $\|v + w\|^2 + \|v - w\|^2 = 2\|v^2\| + 2\|w\|^2$

The proof of these is elementary and may be found in chapter 4 of Rudin [6].

In the infinite-dimensional case, we also must be careful about the kinds of representations we allow. In particular, it is natural to consider representations which arise from a continuous action of a topological group G. We will explain what this means precisely later, but in order to do so, we must establish the notion of bounded operators:

Definition 2 (Operator norm)**.** If X, Y are normed spaces and $T : X \to Y$, then T is called **bounded** if there exists a constant C such that $\|T(x)\| \leq C\|x\|$ for all $x \in X$. The **operator norm** of T, denoted $\|T\|_{op}$, is the infimum of such C. The bounded operators from X to Y are denoted $\mathcal{B}(X, Y)$.

Proposition 2. If X and Y are normed spaces, then a linear operator $T : X \to Y$ is bounded if and only if it is continuous. For a proof, see [6].

Definition 3 (Dual Space)**.** If H is a Hilbert space, then the **dual space** H^* is the set of bounded linear operators $T : H \to \mathbb{C}$.

Theorem 1 (Projection Theorem)**.** Let H be a Hilbert space and $V \leq H$ be a closed subspace. Then $H = V \oplus V^\perp$, where $V^\perp = \{u \in H : \langle u, v \rangle = 0 \ \forall v \in V\}$. For a proof, refer again to chapter 4 of Rudin [6].

Theorem 2 (Frechet-Riesz Representation Theorem)**.** Let H be a Hilbert space with inner product $\langle -, - \rangle$. Then for each $\omega^* \in H^*$ there exists a unique $\omega \in H$ such that $\omega^* = \langle -, \omega \rangle$. Furthermore, $\|\omega^*\|_{op} = \|\omega\|$. This induces an isometric antilinear bijection $I : H \to H^*$ via $I(\omega) = \omega^*$, and we call ω^* the dual vector of ω. For the proof, see Rudin [6].

The Peter-Weyl theorem, as alluded to in the foreword, concerns matrix elements of irreducible representations. In order to define matrix elements in the first place, one needs an orthonormal basis to define matrix coefficients. This is another reason why Hilbert space representations are so central to this theorem.

Definition 4 (Orthonormal Basis)**.** A system $B = \{e_i\}_{i \in I} \subseteq H$ is called an **orthonormal basis** for a Hilbert space H if any $v \in H$ may be written uniquely as a norm-convergent series in B and the elements of B are pairwise-orthonormal, meaning $\langle e_i, e_j \rangle = \delta_{ij}$ for all $i, j \in I$.

Proposition 3 (Existence of Orthonormal Bases)**.** Every Hilbert space possesses an orthonormal basis. If we assume the Axiom of Choice, this holds for all Hilbert spaces, but we can avoid this with the additional assumption that the Hilbert space is separable. An orthonormal system that is dense in H is an orthonormal basis for H. The proof may be found in [6].

Definition 5 (Dual Map)**.** The **dual** A^* of a linear operator $A : X \to Y$ is defined as $A^* : Y^* \to X^*$ via $A^* v^* = v^* \circ A$. Note that if Y is equipped with an inner product, then for every $u \in Y$, the inner product with u defines an element of the dual space $I(u) \equiv \langle -, u \rangle_Y$. Then $A^* I(u) = \langle A(-), u \rangle_Y$.

Definition 6 (Adjoint map)**.** Let X and Y be Hilbert spaces, and consider the maps I_X, I_Y as defined in the Frechet-Riesz representation theorem. For a bounded linear map $A : X \to Y$, we define the **adjoint** $A^\dagger : Y \to X$ such that $A^* \circ I_Y = I_X \circ A^\dagger$, which exists because I_X is invertible. In other words, for any $v \in X$ and $u \in Y$, we have $(A^* I_Y(u))(v) = I_Y(u)(Av) = \langle u, Av \rangle_Y$ and $(A^* I_Y(u))(v) = I_X(A^\dagger u)(v) = \langle A^\dagger u, v \rangle_X$, so this is equivalent to $\langle u, Av \rangle_Y = \langle A^\dagger u, v \rangle_X$.

Definition 7 (Matrix elements)**.** Fixing an orthonormal basis $\{x_i\}$ and $\{y_i\}$ for Hilbert Spaces X and Y, we define the **matrix elements** of A with respect to these bases to be $A_{ij} \equiv \langle Ax_j, y_i\rangle$. Using this definition, $(A^\dagger)_{ij} = \langle A^\dagger y_i, x_j\rangle_X = \langle y_j, Ax_i\rangle_Y = \overline{\langle Ay_i, x_j\rangle_Y} = \overline{A_{ji}}$. This identifies † with the conjugate-transpose.

Definition 8 (Norm Completion)**.** The completion of a normed space X is defined as

$$\hat{X} = \{\text{Cauchy sequences}\}/\{\text{Null sequences}\}$$

The completion satisfies the following properties:

(1) The inclusion $X \hookrightarrow \hat{X}$ is an isometry.

(2) The image of the inclusion of X into $\hat{X}$ is dense in $\hat{X}$.

(3) The completion is unique up to isometry.

Also note that $\hat{X}$ is trivially closed, so using these properties, if $V \subseteq \overline{V} \leq X$ is a subspace of X then $\hat{V}$ is identified with $\overline{V}$, the closure of V in X.

Definition 9 (Completed direct sum)**.** Given a direct sum $W = \bigoplus_n V_n$, we denote the completion of this direct sum $\hat{W} = \widehat{\bigoplus}_n V_n$.

Measure theory preliminaries

Due to the Hilbert space structure of $L^2(G)$, this space pays a distinguished role in functional analysis and physics. As explained in the foreword, a common statement of the Peter-Weyl theorem involves a decomposition of $L^2(G)$ under the action of G by left-translation, i.e. $gf(h) = f(g^{-1}h)$. For the Peter-Weyl theorem, it is crucial to establish the connection between $L^2(G)$ and the continuous functions. Furthermore, the proof of the theorem itself revolves around the notion of integration with respect to a translationally-invariant measure on G. We will establish these notions and some of their properties in this section.

Definition 10 (Measure space)**.** A **measure space** is a topological space X equipped with a $\sigma-$algebra $\mathcal{A}$ and a measure μ. For a precise definition, refer to Jorge's lecture notes [4].

Definition 11 (L^2 inner product)**.** If X is a measure space and $f, g : X \to \mathbb{C}$ are measurable functions, then the L^2**-inner product** between f, g is defined as

$$\langle f, g\rangle_{L^2} = \int_X f(x)\overline{g(x)}d\mu(x)$$

Theorem 3 (Monotone Convergence Theorem)**.** Let $\{f_n\}$ be a monotone increasing sequence of nonnegative measurable functions on X, and $f_n(x) \to f(x)$ as $n \to \infty$ for every $x \in X$. Then $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$. For a proof, see page 21 of [6].

The following theorem is important because it allows us to approximate L^2 functions with continuous functions. This plays a distinguished role in the setting of a compact group, because continuous functions are uniformly continuous. In particular, this will be crucial later on in our use of nets of convolution operators that approximate the identity.

Theorem 4 (Lusin's Theorem)**.** Suppose that X has finite measure, and $f : X \to \mathbb{C}$ is measurable. Then for any $\epsilon > 0$ there exists $g \in C(X)$ such that $\mu(\{x : f(x) \neq g(x)\}) < \epsilon$. For a proof, see page 55 of [6].

Definition 12. Let (X, μ) be a measure space. The vector space $L^2(X, \mu)$ is formally defined as the classes of square-integrable functions $f : G \to \mathbb{C}$ which are equivalent $\mu-$almost everywhere.

Proposition 4. If (X, μ) is a measure space, then $L^2(X, \mu)$ is a Hilbert space with the L^2-inner product. The proof proceeds by verifying completeness of $L^2(G)$ and the inner product axioms, and may be found in [6].

Remark 1. Let X be a compact measure space equipped with a probability measure. Since X is compact, $C(X)$ is included isometrically into $L^2(X)$, and it can be shown that $L^2(X)$ is complete. By Lusin's theorem, $C(X)$ is dense in $L^2(X)$. This shows that $L^2(X)$ is the completion of $C(X)$ with respect to the L^2-norm.

Remark 2. We include $C(X)$ into $L^2(X)$ by identifying a function with a constant sequence, and we frequently identify elements in $L^2(X)$ with a representative measurable, square-integrable function. Usually it is not a problem to identify an element of $L^2(X, \mu)$ with a function f. However, if $f(x) = g(x)$ for $\mu-$a.e. $x \in X$ then $\|f - g\|_{L^2} = 0$. The equivalence classes are necessary for the norm to be positive-definite.

Later on, we will develop the theory of Hilbert-Schmidt operators, which relies heavily on the following lemma:

Lemma 1 (Product Basis Lemma)**.** If X is a $\sigma-$finite measure space and $\{\psi_\alpha\}_{\alpha \in I}$ form an orthonormal basis for $L^2(X)$, then $\{\psi_\alpha \otimes \psi_\beta\}_{\alpha,\beta \in I}$ form an orthonormal basis for $L^2(X \times X)$.

The proof of the above claim rests on Fubini's Theorem, which we state here:

Theorem 5 (Fubini)**.** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be measure spaces. For any $E \subseteq X \times Y$ and $x \in X$, $y \in Y$, define the $x-$slice and $y-$ slice of E respectively to be $E_x = \{y \in Y : (x, y) \in E\}$ and $E^y = \{x \in X : (x, y) \in E\}$. Then $\int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\mu(y)$, and the measure $\mu \otimes \nu$ defined by $\mu \otimes \nu(A) = \int_X \nu(E_x) d\mu(x)$ defines the product measure on $X \times Y$. For any $f \in L^2(X \times Y)$, define $f_x \equiv f\Big|_{\{x\} \times Y}$. Then $f_x \in L^2(Y)$ for μ-a.e. $x \in X$. The proof may be found on page 164 of [6].

Now we proceed with a proof of the Product Basis Lemma, Lem. 1:

Proof. The orthonormality of $\{\psi_\alpha \otimes \psi_\beta\}_{\alpha,\beta \in I}$ follows directly from the definition of the product measure, so it is sufficient to show that this system is dense in $L^2(X \times X, \mu \otimes \mu)$. Let $f \in L^2(X \times X)$ be arbitrary. Since X is $\sigma-$finite, Fubini's theorem tells us that $f_x \in L^2(X, \mu)$ for $\mu-$a.e. $x \in X$. Therefore for such x we can write $f_x(y) = \sum_{\alpha \in I} g_\alpha(x) \psi_\alpha(y)$ which convergence in $L^2(X, \mu)$, where $g_\alpha(x) = \langle f_x, \psi_\alpha \rangle$. Then by Cauchy-Schwarz,

$$\|g_\alpha\|_2^2 = \int_G |\langle f_x, \psi_\alpha \rangle|^2 d\mu(x) \leq \|f_x\|_2^2 < \infty \ (\mu - \text{a.e. } x) \tag{1}$$

This shows that for such x, we can write $g_\alpha(x) \equiv \sum_{\beta \in I} \langle g_\alpha, \psi_\beta \rangle \psi_\beta(x) = \sum_{\beta \in I} c_{\alpha\beta} \psi_\beta(x)$, where $c_{\alpha\beta} = \langle \langle f_x, \psi_\alpha \rangle, \psi_\beta \rangle$. Substituting this into the expression for f_x gives

$$f(x, y) = f_x(y) = \sum_{\alpha \in I} g_\alpha(x) \psi_\alpha(y) = \sum_{\alpha,\beta \in I} c_{\alpha\beta} \psi_\beta(x) \psi_\alpha(y) \tag{2}$$

with convergence in $L^2(X \times X, \mu \otimes \mu)$ since the expansion holds for $\mu-$a.e. $x, y \in X$. $\square$

Theorem 6. Let G be a locally compact topological group. Then G admits a left-translation invariant measure μ_L and a right-translation invariant measure μ_R. If G is compact, then these two measures agree, and are referred to as the Haar measure [4]. Since the Haar measure is Borel by construction, it takes a finite value on G, and so we will take it to be normalized so that $\mu(G) = 1$.

Remark 3. The Haar measure on a compact group is also invariant under inversion, but we will not use this property.

Lemma 2 (U-substitution)**.** If G is a compact topological group equipped with the Haar measure μ and f is a measurable function on G, then by the invariance of the Haar measure under translations, for any $h, k \in G$ the following holds:

$$\int_G f(x)d\mu(x) = \int_G f(kx)d\mu(x) = \int_G f(xh)d\mu(x)$$

Representation theory preliminaries

In this section, we will establish the groundwork of representation theory. In particular, this section shows one reason that the Haar measure is so important: all finite-dimensional representations are unitarizable and therefore completely reducible. This will be an essential ingredient in establishing the relationship between $C_{\text{alg}}(G)$ and $C(G)^{\text{fin}}$.

Definition 13 (Continuous Group Representation)**.** If G is a topological group, a continuous linear representation of a group G is a vector space V and a group homomorphism $\rho : G \to \mathcal{B}(V)$, the bounded linear operators on V, such that for each $v \in V$ the map $g \mapsto \rho(g)v$ is norm-continuous. *Going forward, we will say "representation" to mean continuous, linear representation.*

Remark 4. While not essential for the rest of the proof, we note that equivalent conditions include

- ρ is continuous when $\mathcal{B}(V)$ is given the strong-operator topology
- $(g, v) \mapsto \rho(g)v$ is continuous in the product topology.

Convergence of T_n to $T \in B(V)$ happens in the **strong-operator** topology on $B(V)$ if $T_n v \to Tv$ for all $v \in V$.

Remark 5. One might wonder how a representation of a topological group could be non-continuous. Such examples exist and are far from pathological. For example, onsider the action of $\mathbb{R}$ on $C(\mathbb{R})$, where a limit of continuous functions may not be continusous. For a compact example, see [10], pages 149-150 for a non-continuous representation of $\mathrm{SO}(n, \mathbb{R})$.

Frequently, either V or ρ alone are referred to as the representation of G if the other is clear from context. We will restrict our attention to vector spaces over $\mathbb{C}$. The Peter-Weyl theorem is concerned with representations of a compact topological group G on Hilbert spaces, so we will only address infinite-dimensional representations when they occur in pre-Hilbert and Hilbert spaces.

Definition 14 (Homomorphism of representations)**.** If (ρ^U, U), (ρ^V, V) are representations of G, then a linear map $\psi : U \to V$ is said to be a **homomorphism of representations** if for all $g \in G$ and $u \in U$, one has $\psi(\rho^U(g)u) = \rho^V(g)\psi(u)$. Such a map is also known as an intertwiner and is called $G-$equivariant. If in addition ψ is invertible, then it is said to be an **isomorphism of representations**.

Definition 15 (Subrepresentations and irreducible representations)**.** If (ρ, V) is a representation of G and $U \leq V$ is a subspace, then U is called G-invariant if $\rho(g)u \in U$ for all $g \in G, u \in U$. If U is a G-invariant subspace, it is a **subrepresentation** with the homomorphism $\rho^U : g \mapsto \rho(g)\Big|_U$. If U itself has no proper subrepresentations, then U is called **irreducible**. We will denote by $\widehat{G}$ the set of isomorphism classes of finite-dimensional irreducible representations of G.

Theorem 7 (Schur's Lemma)**.** Let (ρ^V, V) and (ρ^U, U) be irreducible representations of G over $\mathbb{C}$, and let $\psi : V \to U$ be an homomorphism of representations. Then ψ is either zero or invertible. In particular, if V is finite-dimensional, then ψ acts on V by the constant $\psi = \mathrm{tr}(\psi)/\dim V$.

Proof. We will first prove that $\ker(\psi)$ is a subrepresentation of V. If $v \in \ker(\psi)$ and $g \in G$ is arbitrary, then $\psi(\rho^V(g)v) = \rho^U(g)\psi(v) = 0$, which shows that $\rho^V(g)v \in \ker(\psi)$. Since V is irreducible, it has no proper $G-$invariant subspaces, so either $\ker(\psi) = V$ or $\ker(\psi) = \{0\}$.

Now we will prove that $\operatorname{im}(\psi) \leq U$ is a subrepresentation. Let $w \in \operatorname{im}(\psi)$. Then there exists $v \in V$ such that $\psi(v) = w$, and $\rho^U(g)\psi(v) = \psi(\rho^V(g)v)$, so $\rho^U(g)\psi(v) \in \operatorname{im}(\psi)$. This shows that $\operatorname{im}(\psi) = U$ or $\operatorname{im}(\psi) = \{0\}$. In the first case, ψ is invertible, and in the second case, ψ is zero.

Now assume V is finite-dimensional. Since $\mathbb{C}$ is algebraically closed, ψ must have at least one eigenvalue λ. Thus $\psi - \lambda \operatorname{Id}$ is a morphism of representations that is not invertible, so by the above argument, it must be zero. This sets $\psi = \lambda \operatorname{Id}$. Taking the trace on both sides, $\operatorname{tr}(\psi) = \lambda \operatorname{tr}(\operatorname{Id}) = \lambda \dim V$. Solving for λ proves the theorem. ☐

Remark 6. Schur's lemma shows why we work over $\mathbb{C}$, an algebraically closed field, and why the same proof would not work over a non-algebraically-closed field such as $\mathbb{R}$. Namely, since ψ in the proof above may not have an eigenvalue over $\mathbb{R}$, Schur's Lemma can only say that ψ in the above proof must be zero or invertible, and need not act as a scalar.

Definition 16 (Unitary Representations)**.** If (ρ, V) is a representation of G and $\langle -, - \rangle$ is an inner product on V, then the representation is called **unitary** if for all $g \in G$, $\rho(g)$ is unitary with respect to this inner product. In other words, the following holds: for any $u, v \in V$, $\langle \rho(g)u, \rho(g)v \rangle = \langle u, v \rangle$. In this case, the inner product is said to be $G-$invariant. This is equivalent to $\rho(g)^\dagger = \rho(g^{-1})$.

Proposition 5 (Unitarizability of finite-dimensional G-reps)**.** Let G be a compact topological group. Then any finite-dimensional representation has a $G-$invariant inner product. Such a representation is called unitarizable.

Proof. Let μ be the normalized Haar measure on G. Let (ρ, V) be a finite-dimensional representation of G. If $(-, -)$ is any inner product on V, then we can construct a bilinear form by averaging:

$$\langle u, v \rangle \equiv \int_G (\rho(g)u, \rho(g)v) d\mu(g) \tag{3}$$

for any $u, v \in V$. It is quick to check that this defines an inner product. Using the $u-$substitution $g \to h^{-1}g$, we find

$$\begin{aligned} \langle \rho(h)u, \rho(h)v \rangle &= \int_G (\rho(hg)u, \rho(hg)v) d\mu(g) && (4) \\ &= \int_G (\rho(g)u, \rho(g)v) d\mu(g) && (5) \\ &= \langle u, v \rangle && (6) \end{aligned}$$

This shows that $\langle -, - \rangle$ is $G-$invariant. ☐

Remark 7. Averaging an inner product is known as the Weyl averaging trick.

Proposition 6 (Uniqueness of $G-$invariant inner products)**.** Let (ρ, V) be a finite-dimensional irreducible representation of G. Then any two $G-$invariant inner products on V are related by a positive constant.

Proof. Define the linear map $\Phi : V \to V$ via $\langle \Phi(u), v \rangle_2 = \langle u, v \rangle_1$ for all $v \in V$. Then

$$\langle \Phi(\rho(g)u), v \rangle_2 = \langle \rho(g)u, v \rangle_1 = \langle u, \rho(g^{-1})v \rangle_1 = \langle \Phi(u), \rho(g^{-1})v \rangle_2 = \langle \rho(g)\Phi(u), v \rangle$$

Since this holds for any v, this shows that Φ is a homomorphism of representations. By Schur's lemma, $\Phi(u) = \lambda u$ for some scalar λ, so $\langle u, v \rangle_1 = \lambda \langle u, v \rangle_2$. In particular, $\langle v, v \rangle_1 = \lambda \langle v, v \rangle_2$, so λ is real and positive. ☐

Remark 8. The above proposition is important because it implies that a basis for V which is orthonormal with respect to some $G-$invariant inner product is orthogonal with respect to all $G-$invariant inner products.

Definition 17 (Complete reducibility)**.** A representation (ρ, V) of G is called **completely reducible** or semisimple if for every $G-$invariant subspace $W \leq V$, then there is another $G-$invariant subspace U such that $V = W \oplus U$.

Theorem 8. Every unitarizable representation of G over $\mathbb{C}$ is completely reducible. In particular, all finite-dimensional representations of a compact group G are completely reducible.

Proof. Suppose $U \leq V$ is a $G-$invariant subspace. Since V is unitarizable, take $\langle -, - \rangle$ to be a $G-$invariant inner product. Let $v \in U$ and $w \in U^\perp$ be arbitrary. Then for any $g \in G$, $\langle v, \rho(g)w \rangle = \langle \rho(g^{-1})v, w \rangle = 0$, where we have used the fact that $\rho(g^{-1})v \in U \perp w$. This shows that $\rho(g)w \in U^\perp$, so $V = U \oplus U^\perp$ where $U^\perp$ is $G-$invariant. $\square$

Definition 18 (Contragredient representation)**.** A representation (ρ^V, V) induces a representation in the dual space (ρ^{V^*}, V^*) called the **contragredient** representation. This representation is defined by $\rho^{V^*}(g)v^* = \rho^V(g^{-1})^* v^*$. We can see from this definition that V^* is irreducible iff V is.

Remark 9. If the representation is unitary, this means that $\rho^{V^*}(g)\langle -, v \rangle = \langle \rho^V(g^{-1})(-), v \rangle = \langle -, \rho^V(g)v \rangle$.

Definition 19 (Left-regular representation)**.** Given a group G, the left-regular representation of G is $(L, L^2(G))$, where $L^2(G)$ is the vector space of square-integrable functions and L_g acts on $f \in L^2(G)$ by left-translation: $L_g f(h) = f(g^{-1}h)$.

Remark 10. In the case of finite groups, the left-regular representation is the group-ring $\mathbb{C}[G]$. The action of G on $\mathbb{C}[G]$ is naturally defined via left-multiplication, i.e., if $a = \sum_{g \in G} a_g g \in \mathbb{C}[G]$, then $\rho(h)a = \sum_{g \in G} a_g hg$. If we write $f_a(g) \in \mathrm{Fun}(G \to \mathbb{C})$ such that $f_a(g) = a_g$, then an element of $\mathbb{C}[G]$ is naturally identified with an element of $\mathrm{Fun}(G \to \mathbb{C})$. Reindexing the sum, we find $\rho(h)g = \sum_{h^{-1}g \in G} f_a(h^{-1}g)g$, so $\rho(h)a$ is identified with $f_{\rho(h)a}(g) = f_a(h^{-1}g) = L_h f_a(g)$.

Matrix coefficients

Definition 20 (Representative function)**.** If V is a representation of G, then the subspace of vectors $v \in V$ which generate finite-dimensional subspaces under the action of G is denoted V^{fin}. The subspace $C(G)^{\text{fin}}$ under the action of G by left-translation are called representative functions.

Definition 21 (Matrix coefficient)**.** If (ρ^V, V) is any finite-dimensional representation of G, and $w \in V^*$, $v \in V$, then $\rho^V_{w^*,v} : G \to \mathbb{C}$ via

$$\rho^V_{w^*,v}(g) = \langle \rho^V(g)v, w \rangle = w^*(\rho^V(g)v)$$

is a continuous function called a matrix coefficient. The set of matrix coefficients is denoted $C_{\text{alg}}(G)$.

Remark 11. If V is a finite-dimensional representation, we will frequently fix an orthonormal basis for V with respect to a $G-$invariant inner product $\{v_i\}$ and write $\rho^V_{v_i^*,v_j}(g) = \rho^V_{ij}(g)$ for convenience.

Example 1. For an example of matrix elements of a finite group, consider $D_{2n} = \langle a, x | a^n = x^2 = e, axa = a^{-1} \rangle$, where n is odd. An irreducible 2D representation of D_{2n} is the following:

$$a \mapsto \begin{pmatrix} e^{2\pi i/n} & 0 \\ 0 & e^{-2\pi i/n} \end{pmatrix} \qquad x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7}$$

Taking the upper-left matrix element, we get a function $D_{2n} \mapsto \mathbb{C}$ via

$$a^k \mapsto e^{2\pi i k/n} \qquad a^k x \mapsto 0 \tag{8}$$

Example 2. Another example that will be crucial for Fourier analysis is the representation theory of the unit circle S^1 over $\mathbb{C}$. This is a Lie group, and being Abelian, all representations are one-dimensional. All irreducible representations, indexed by n, are given by $\rho_n : \theta \mapsto e^{2\pi i n\theta}$. These are matrix elements, and we recognize them as the familiar Fourier series.

Lemma 3. Let (ρ^U, U) and (ρ^V, V) be finite-dimensional representations of G, and let $T : U \to V$ be a linear map. Then

$$\widetilde{T} \equiv \int_G \rho^U(g) T \rho^V(g^{-1}) d\mu(g) = \frac{\text{tr}\{T\}}{\dim V} \delta_{UV}$$

Where $\delta_{UV} = 1$ if $U \cong V$ and zero otherwise.

Proof. We will show that $\widetilde{T}$ is an intertwiner. Using a u-substitution,

$$\rho^U(h)\widetilde{T} = \int_G \rho^U(hg) T \rho^V(g^{-1}) d\mu(g) \tag{9}$$

$$= \int_G \rho^U(g) T \rho^V(g^{-1}h) d\mu(g) \tag{10}$$

$$= \widetilde{T}\rho^V(h) \tag{11}$$

This shows that $\widetilde{T}$ is an intertwiner. By Schur's lemma, if U and U are non-isomoprhic, $\widetilde{T} = 0$. Otherwise, $V \cong U$. Note that if $\psi : V \to U$ is an isomorphism of irreducible representations, then by definition, $\rho^U(g)\psi = \psi\rho^V(g)$. Since ψ acts by a scalar due to Schur's Lemma, $\rho^V = \rho^U$. Since the representation is unitary, cyclicity of the trace shows that

$$\widetilde{T} = \frac{1}{\dim V} \text{tr}\bigg\{ \int_G \rho^V(g) T \rho^V(g^{-1}) \bigg\} d\mu(g) \tag{12}$$

$$= \frac{1}{\dim V} \int_G \text{tr}\{\rho^V(g) T \rho^V(g^{-1})\} d\mu(g) \tag{13}$$

$$= \frac{1}{\dim V} \int_G \text{tr}\{T\} d\mu(g) \tag{14}$$

$$= \frac{\text{tr}\{T\}}{\dim V} \tag{15}$$

This proves the claim. □

Proposition 7 (Schur orthogonality of matrix coefficients)**.** Let (ρ^U, U) and (ρ^V, V) be finite-dimensional irreducible unitary representations. Fix orthonormal bases $\{e_i\}$ for U and $\{d_i\}$ for V with respect to a $G-$invariant inner product, with $e_i = d_i$ if $U = V$. Then

$$\langle \rho^V_{ij}, \rho^U_{kl} \rangle_{L^2} = \frac{\delta_{UV}\delta_{ik}\delta_{jl}}{\dim V}$$

In other words, the matrix coefficients form an orthonormal system with respect to this dilated inner product.

Proof. Let $\{e_i^*\}$ and $\{d_i^*\}$ be dual bases for U^* and V^* respectively. Since the bases are chosen with respect to a $G-$invariant inner product, $\rho^V(g)^\dagger = \rho^V(g^{-1})$. By definition, $\rho^U_{kl}(g) = e_k^* \rho^U(g) e_l$ and $\rho^V_{ij}(g) = d_i^* \rho^V(g) d_j$. Expanding, we see

$$\langle \rho^V_{ij}, \rho^U_{kl} \rangle_{L^2} = \int_G (d_i^* \rho^V(g) d_j)(\overline{e_k^* \rho^U(g) e_l}) dg \tag{16}$$

$$= \int_G (d_i^* \rho^V(g) d_j)(e_l^* \rho^U(g)^\dagger e_k) dg \tag{17}$$

$$= d_i^* \bigg(\int_G \rho^V(g)(d_j e_l^*) \rho^U(g^{-1}) dg \bigg) e_k \tag{18}$$

Since $d_j e_l^* : U \to V$, we apply the previous lemma to it. If $U \neq V$, then the term in parenthesis is zero and the entire expression vanishes. If $U = V$, then $d_j = e_j$ and applying the previous lemma,

$$= e_i^* \left(\int_G \rho^V(g)(e_j e_l^*)\rho^U(g^{-1})dg \right) e_k \tag{19}$$

$$= \frac{\operatorname{tr}\{e_j e_l^*\}}{\dim V} e_i^* e_k \tag{20}$$

$$= \frac{\delta_{jl}\delta_{ik}}{\dim V} \tag{21}$$

Putting these two cases together establishes the claim. □

Remark 12. The chosen bases e_i, d_i to express the matrix coefficients needed to be orthonormal with respect to a $G-$invariant inner product in order to use $\rho(g^{-1})_{ij} = \overline{\rho(g)_{ji}}$. In spite of this fact, we still arrived at an orthonormal system with respect to the standard inner product on $L^2(G)$, which will be indispensable for Fourier analysis.

Proposition 8. $C_{\text{alg}}(G)$ is a self-adjoint unital subalgebra of $C(G)$ under pointwise sum and product.

Proof. Let $(V, \rho^V), (W, \rho^W)$ be finite-dimensional representations of G. Let $\omega_1, v_1 \in V$ be arbitrary. We observe that $C_{\text{alg}}(G)$ is closed under scalar multiplication:

$$\lambda \rho^V_{\omega_1^*, v_1} = \rho^V_{\omega_1^*, \lambda v_1} \tag{22}$$

Now let $\omega_2, v_2 \in W$. Then we can show closure under pointwise sum:

$$\rho^{V\oplus W}_{(\omega_1 \oplus \omega_2)^*, v_1 \oplus v_2}(g) = (\omega_1^* \oplus \omega_2^*)(\rho^V(g)v_1 \oplus \rho^W(g)v_2) \tag{23}$$

$$= \omega_1^*(\rho^V(g)v_1) + \omega_2^*(\rho^V(g)v_1) + \omega_1^*(\rho^W(g)v_2) + \omega_2^*(\rho^W(g)v_2) \tag{24}$$

$$= \omega_1^*(\rho^V(g)v_1) + \omega_2^*(\rho^W(g)v_2) \tag{25}$$

$$= \rho^V_{w_1^*, v_1}(g) + \rho^W_{w_2^*, v_2}(g) \tag{26}$$

where we have used the notation $v_1 \oplus v_2$ to denote the external direct sum, emphasizing the fact that $\omega_1\Big|_W = \omega_2\Big|_V = 0$ even when $W = V$. We can establish a similar identity for multiplication using the tensor product:

$$\rho^{V\otimes W}_{(\omega_1 \otimes \omega_2)^*, v_1 \otimes v_2}(g) = (\omega_1^* \otimes \omega_2^*)(\rho^V(g)v_1 \otimes \rho^W(g)v_2) \tag{27}$$

$$= \omega_1^*(\rho^V(g)v_1)\omega_2^*(\rho^W(g)v_2) \tag{28}$$

$$= \rho^V_{\omega_1^*, v_1}(g)\rho^W_{\omega_2^*, v_2}(g) \tag{29}$$

Where we have used the definition of the tensor product representation and the identification $(\omega_1 \otimes \omega_2)^* = \omega_1^* \otimes \omega_2^*$ in finite dimensions.

Finally, to prove that this algebra is self-adjoint, we examine the contragredient representation. Since the representation is unitary, we have $\rho^{V^*}(g)w^* = \langle -, \rho^V(g)w \rangle_V = (\rho^V(g)w)^*$. Since the map $v \mapsto v^*$ is antilinear, it preserves the inner product on V^* defined as $\langle v^*, w^* \rangle_{V^*} = \overline{\langle v, w \rangle_V}$. This gives

$$\overline{\rho^V_{w^*, v}(g)} = \overline{\langle \rho^V(g)v, w \rangle_V} = \langle (\rho^V(g)v)^*, w^* \rangle_{V^*} = \langle \rho^{V^*}(g)v^*, w^* \rangle_{V^*} = \rho^{V^*}_{w, v^*}(g) \tag{30}$$

This shows that $C_{\text{alg}}(G)$ is self-adjoint. □

Proposition 9. For each $V \in \widehat{G}$, fix a $G-$invariant inner product and an orthonormal basis $\{\psi_i^V\}_{i=1}^{\dim V}$. Then $C_{\text{alg}}(G)$ is equal to the set of finite linear combinations of $\{\rho_{ij}^V\}_{i,j,V \in \widehat{G}}^{\dim V}$.

Proof. We have shown above that $C_{\text{alg}}(G)$ is closed under linear combinations, and thus contains $\text{span}_{V\in\widehat{G},i,j}\{\rho_{ij}^V\}$, so we must show containment in the other direction. Suppose that $\rho_{\omega^*,\eta}^U \in C_{\text{alg}}(G)$ is arbitrary. Since U is finite-dimensional and thus completely reducible, we can write $U = \bigoplus_{V\in\widehat{G}} V^{\oplus n_V}$ where n_V is an integer representing the multiplicity of V in U and is only greater than zero for finitely many V. By definition, we can decompose ω and η as $\omega = \sum_{V\in\widehat{G}}\sum_{i=1}^{n_V}\omega_i^V$ and $\eta = \sum_{V\in\widehat{G}}\sum_{j=1}^{n_V}\eta_j^V$, where $\omega_i^V, \eta_j^V \in V$ for all i, j. Thus we have

$$\rho_{\omega^*,\eta}^U = \sum_{V\in\widehat{G}}\sum_{i,j=1}^{n_V}\rho_{(\omega_i^V)^*,\eta_j^V}^V$$

Expanding $\omega_i^V = \sum_{k=1}^{\dim V}\alpha_{ik}^V\psi_k^V$ and $\eta_j^V = \sum_{l=1}^{\dim V}\beta_{jl}^V\psi_l^V$, we can express $\rho_{(\omega_i^V)^*,\eta_j^V}^V = \sum_{k,l=1}^{\dim V}\overline{\alpha_{ik}^V}\beta_{jl}^V\rho_{kl}^V$. Thus $\text{span}\{\rho_{ij}^V\}_{i,j,V\in\widehat{G}}^{\dim V} = C_{\text{alg}}(G)$ as claimed. □

Proposition 10. If $C(G)^{\text{fin}}$ is considered as a representation of G by left-translation, then $f \in C(G)^{\text{fin}}$ is a (finite) linear combination of matrix coefficients.

Proof. Let $f \in V^{\text{fin}}$ be arbitrary. Then $f \in U \leq V$, where U is a representation of G with dimension N. Fix an orthonormal basis $\{v_i\}_{i=1}^N$ for U. Writing L_g as a matrix in this basis defines a representation $\rho(g)$. Since f is a linear combination of the $\{v_i\}$, it is enough to prove the claim for v_i. We see that

$$v_i(g) = L_{g^{-1}}v_i(e) = [\rho(g^{-1})v_i](e) = \sum_j \rho(g^{-1})_{ji}v_j(e) = \sum_i v_j(e)\overline{\rho_{ij}(g)} \tag{31}$$

Since $\overline{\rho_{ij}} \in C_{\text{alg}}(G)$ this expresses v_i as a linear combination of matrix coefficients. □

Lemma 4. $C_{\text{alg}}(G) = C(G)^{\text{fin}}$.

Proof. We have already shown that $C(G)^{\text{fin}} \leq C_{\text{alg}}(G)$ so we need to show that $\rho_{\omega^*,v}^V \in C_{\text{alg}}(G)$ for arbitrary $V \in \widehat{G}$ generates a finite-dimensional subspace. Fix a basis $\{v_i\}$ for V. We claim that $\text{span}\{\rho_{v_i^*,v}\}_{i=1}^{\dim V}$ contains all orbits of $\rho_{\omega^*,v}^V$ under left-translation. We observe that

$$\begin{aligned}
L_g\rho_{\omega^*,v}^V(h) &= \langle\rho^V(g^{-1}h)v, \omega\rangle && (32)\\
&= \langle\rho^V(h)v, \rho^V(g)\omega\rangle && (33)\\
&= \langle\rho^V(h)v, \sum_{ij}\rho_{ij}^V(g)\omega_i v_j\rangle && (34)\\
&= \sum_{ij}\overline{\rho_{ij}^V(g)\omega_i}\langle\rho^V(h)v, v_j\rangle && (35)\\
&= \sum_{ij}\overline{\rho_{ij}^V(g)\omega_i}\rho_{v_j^*,v}^V(h) && (36)
\end{aligned}$$

Thus $L_g\rho_{\omega^*,v}^V \in \text{span}\{\rho_{v_i^*,v}^V\}_{i=1}^{\dim V}$ which has dimension $\dim V$, so $\rho_{\omega^*,v}^V$ generates a finite-dimensional subspace of $C(G)$. This shows that $C(G)^{\text{fin}} = C_{\text{alg}}(G)$. □

Spectral theorem for compact self-adjoint operators

In order to prove the theorem, we need to be able to "carve out" finite-dimensional left-invariant subspaces of $L^2(G)$. It turns out that a natural way to do this is by defining a compact operator that commutes with the action of G. In this section, we will show how a compact, self-adjoint operator gives us many orthogonal finite-dimensional subspaces of $L^2(G)$.

Lemma 5. Let T be a self-adjoint operator on a Hilbert space H. Then $\|T\|_{op} = \sup_{|x|=1}|\langle Tx, x\rangle|$.

Proof. Let $M = \sup_{\|x\|=1} \langle Tx, x\rangle$. For any $\|x\| = 1$, By Cauchy-Schwarz, $|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|_{op}$. Assume that $Tx \neq 0$ and let $y = Tx/\|Tx\|$. Since T is self-adjoint, we have $\langle Tx, y\rangle = \|Tx\| = \langle x, Ty\rangle$. Thus

$$\langle T(x+y), x+y\rangle = \langle Tx, x\rangle + 2\|Tx\| + \langle Ty, y\rangle \tag{37}$$
$$\langle T(x-y), x-y\rangle = \langle Tx, x\rangle - 2\|Tx\| + \langle Ty, y\rangle \tag{38}$$

Subtracting these two equations,

$$4\|Tx\| = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \tag{39}$$
$$\leq M\|x+y\|^2 + M\|x-y\|^2 = 4M \tag{40}$$

where in the last step we have used the parallelogram identity. □

Definition 22 (Compact Operator)**.** An operator $T : H \to H$, where H is a Hilbert space, is called compact if $T\overline{\mathbb{B}}^1$ is precompact in H.

Lemma 6. Let T be a compact self-adjoint operator on a Hilbert space H. Then T has an eigenvalue of magnitude $\|T\|_{op}$.

Proof. Choose a maximizing sequence of unit vectors x_i for $|\langle Tx_i, x_i\rangle|$. Since $\langle Tx_i, x_i\rangle = \overline{\langle Tx_i, x_i\rangle}$, take $\langle Tx_i, x_i\rangle \to \lambda = \pm\|T\|_{op}$. Since T is compact, reduce to a subsequence y_i such that $Ty_i \to v$ for some $v \in H$. We observe that

$$|\langle Ty_i, y_i\rangle| \leq |Ty_i| \leq \|T\|_{op} = |\lambda|$$

This shows that $\|Ty_i\| \to |\lambda|$. Thus we see that

$$\|\lambda y_i - Ty_i\|^2 = \lambda^2\|y_i\|^2 + \|Ty_i\|^2 - 2\lambda\langle Ty_i, y_i\rangle \to 0$$

Since $Ty_i \to v$, $\|\lambda y_i - v\| \leq \|\lambda y_i - Ty_i\| + \|Ty_i - v\| \to 0$, so $\lambda y_i \to v$. Therefore $y_i \to v/\lambda$, and by continuity of T, $T(v/\lambda) = v$, so $Tv = \lambda v$ by linearity. This shows that v is the required eigenvector of T with eigenvalue $\lambda = \pm\|T\|_{op}$. □

Theorem 9 (Spectral theorem for compact self-adjoint operators)**.** If $T : H \to H$ is a compact self-adjoint operator on a Hilbert space H, then

$$S = \ker(T) \oplus \bigoplus_{i=1}^{\infty} E_{\lambda_i}$$

is dense in H, where E_λ is the eigenspace of T with eigenvalue λ. T has at most countably many eigenvalues λ_i satisfying $\lambda_i \to 0$ and all eigenspaces are finite-dimensional.

Proof. Let E be the set of all nonzero eigenvalues, and suppose E is uncountable. Then there exists ξ such that $\{\lambda \in E : |\lambda| > \xi\}$ is uncountable, otherwise we could form E from a countable union of countable sets. T has orthogonal eigenvectors, so $\{Tx : x \in E_\lambda, \|x\| = 1, |\lambda| > \xi\}$ is not Cauchy, and therefore has no convergent subsequences. This contradicts the compactness of T. This shows that T has countably many nonzero eigenvalues. Similarly, if λ_i does not approach zero, then there is a sequence v_i with $v_i \in E_{\lambda_i}$ and $\|v_i\| = 1$. Since all the v_i are pairwise orthogonal, this sequence is not Cauchy, and so it has no convergent subsequence, again contradicting the compactness of T. This shows that $\lambda_i \to 0$ as $i \to \infty$. Lastly, let $\overline{S}$ be the closure of S. Let $S^\perp$ be the orthogonal complement of $\overline{S}$. If $S^\perp \neq \{0\}$, then by the previous lemma, $T\Big|_{S^\perp}$ has a nonzero eigenvalue, which contradicts the construction of S. This shows that $S^\perp = \{0\}$, proving that S is dense in H. □

In the next section, we will develop the theory of Hilbert-Schmidt operators, for which it will be useful to have another characterization of compact operators:

Lemma 7. If $T : L^2(X) \to L^2(X)$ is the uniform-norm limit of finite-rank operators, then T is compact.

Proof. Suppose that $T = \lim_{n\to\infty} T_n$ with convergence in the operator norm, where T_n is finite rank. Let $\{x_n\}$ be a bounded sequence. Since the image of T_n is finite-dimensional and T_n is a bounded operator, Bolzano-Weierstrass holds that T_n is compact. Let $x_n^{(1)}$ be the subsequence of x_n such that $T_1 x_n$ converges. Then inductively, suppose that $x_n^{(k)}$ is constructed such that $T_k x_n^{(k)}$ converges. Then put $x_n^{(k+1)}$ to be the subsequence such that $T_{k+1} x_n^k$ converges. Lastly, let $y_n = x_n^{(n)}$, and we will show that Ty_n is Cauchy. For any $\epsilon > 0$, there is an N such that for all $n, m > N$ we have $d(T_k y_n, T_k y_m) < \epsilon$. By the definition of uniform-norm convergence, there exists M such that $d(Ty_n, T_k y_n) < \epsilon$ for all $k > M$. Thus $d(Ty_n, Ty_m) \leq d(Ty_n, T_k y_n) + d(T_k y_n, T_k y_m) + d(Ty_m, T_k y_m) < 3\epsilon$. This proves that Ty_n is Cauchy, and since $L^2(X)$ is complete, Ty_n converges, so T is compact. $\square$

Remark 13. The converse holds as well for a Hilbert space.

Remark 14. Note that the above statement is emphatically not true for convergence in the strong topology. In fact, finite-rank operators are dense in the space of operators with respect to the strong topology.

Hilbert-Schmidt operators

Building on the previous section, if we can construct compact operators that commute with the action of G, we will obtain a decomposition of $L^2(G)$ into invariant eigenspaces of this operator, of which all the non-zero eigenspaces will be finite-dimensional. An easy way to construct such operators is through integral operators. It turns out that many of these integral operators are compact, which will be established in this section via the Hilbert-Schmidt Theorem.

Definition 23 (Hilbert-Schmidt Operator)**.** A bounded linear operator $T : L^2(X, \mu) \to L^2(X, \mu)$ is called a **Hilbert-Schmidt** operator if it is representable as

$$(Tf)(x) = \int_X K(x, y) f(y) d\mu(y)$$

where $K \in L^2(X \times X, \mu \otimes \mu)$ is called the Hilbert-Schmidt kernel of T.

Remark 15. The Hilbert-Schmidt operators, in many ways, naturally generalize the operators on finite-dimensional spaces. Multiplication of a vector f by a matrix K is $(Kf)_i = \sum_j K_{ij} f_j$. Integral operators generalize this notion with $K(x, y)$ and $f(y)$.

Theorem 10 (Hilbert-Schmidt)**.** Every Hilbert-Schmidt operator T is compact. *This appears as exercise 3.2 in Bump (sadly the bump function is not eponymous) [2].*

Proof. We will show that T is the uniform limit of finite-rank operators. Due to the lemma above, we can take an orthonormal basis $\{\psi_\alpha\}_{\alpha\in I}$ for a possibly uncountable indexing set I. We can decompose the kernel K into $K(x, y) = \sum_{\alpha,\beta\in I} c_{\alpha\beta} \psi_\alpha(x) \psi_\beta(y)$, with converges in the strong-operator topology with respect to the L^2-norm. Note that

$$\|K\|_{L^2(X\times X)} = \sum_{\alpha,\beta} |c_{\alpha\beta}|^2 < \infty$$

due to the fact that K is a Hilbert-Schmidt kernel. Since an uncountable sequence of nonzero real numbers cannot converge, only countably many $c_{\alpha\beta}$ are nonzero. Then for any $f \in L^2(X)$ we can write

$$(Tf)(x) = \int_X K(x,y)f(y)dy \tag{41}$$
$$= \int_X \sum_{i,j=1}^{\infty} c_{ij}\psi_i(x)\psi_j(y)f(y)dy \tag{42}$$
$$= \sum_{i,j=1}^{\infty} c_{ij}\psi_i(x)\int_X \psi_j(y)f(y)dy \tag{43}$$
$$= \sum_{i,j=1}^{\infty} c_{ij}\langle f, \overline{\psi_j}\rangle\psi_i(x) \tag{44}$$

where we have applied the Monotone Convergence Theorem in the third step. Each of the summands $f \mapsto \langle\overline{\psi_\beta}, f\rangle$ is a rank-1 operator. Therefore, let

$$T_n = \sum_{i,j<n} c_{ij}\langle -, \overline{\psi_j}\rangle\psi_i$$

Convergence is guaranteed in the strong-operator norm, so we must show uniform-norm convergence. Note that by construction, $T - T_n = \sum_{i,j>n} c_{ij}\langle -, \overline{\psi_j}\rangle\psi_i$. Let $f \in L^2(X)$ be arbitrary, and decompose $f = \sum_{\gamma\in I} b_\gamma\overline{\psi_\gamma}$. Then

$$\|(T-T_n)f\|_2^2 = \left\|\sum_{i,j\geq n} c_{ij}\langle\sum_{\gamma\in I}\overline{\psi_\gamma}b_\gamma, \overline{\psi_j}\rangle\psi_i\right\|_2^2 = |\sum_{i,j\geq n} c_{ij}b_j| \leq \|f\|_2^2\sum_{i,j\geq n}|c_{ij}|^2 \to 0$$

As shown above, $\|K_{L^2(X\times X)}\| = \sum_{i,j}|c_{ij}|^2 < \infty$, so the tail of the sum has to converge to zero. This gives T as a uniform limit of finite-rank operators. ☐

Remark 16. There is an alternative approach to Peter-Weyl using the Arzela-Ascoli theorem, which requires the Hilbert-Schmidt kernel to be uniformly continuous. The compactness assumption is used to apply this result to all continuous kernels. The Hilbert-Schmidt theorem shows that this holds more generally.

Lemma 8. If $K(x,y) = \overline{K(y,x)}$, then T is self-adjoint.

Proof.

$$\langle Tf_1, f_2\rangle = \int_X\left(\int_X K(x,y)f_1(y)dy\right)\overline{f_2(x)}dx \tag{45}$$
$$= \int_X f_1(y)\overline{\int_X K(y,x)f_2(x)dxdy} \tag{46}$$
$$= \langle f_1, Tf_2\rangle \tag{47}$$

☐

Definition 24 (Convolution)**.** Let T_ϕ be the Hilbert-Schmidt operator with kernel $K(x,y) = \phi(y^{-1}x)$, where $\phi \in L^2(G)$. We refer to $T_\phi f \equiv \phi * f$ as the **convolution** of ϕ with f. Furthermore, we see that T_ϕ is self-adjoint if $\phi(g^{-1}) = \overline{\phi(g)}$.

Lemma 9. The eigenspaces of T_ϕ are left-translation invariant.

Proof. We only need to show that T_ϕ and L_g commute as operators, which will imply that L_g preserves the eigenspaces of T_ϕ. Using a u-substitution,

$$T_\phi(L_g f)(x) = \int_G \phi(h^{-1}x) f(g^{-1}h) dh \tag{48}$$

$$= \int_G \phi(h^{-1}[g^{-1}x]) f(h) dh \tag{49}$$

$$= L_g(T_\phi f)(x) \tag{50}$$

□

Remark 17. The convolution operator makes $C(G)$, and indeed $L^2(G)$ when G has finite measure, into an algebra. This algebra is natural to consider for finite groups under the identification of $\text{Fun}(G \to \mathbb{C})$ with $\mathbb{C}[G]$ via $a \mapsto f_a$ as described previously. Under this identification, for any $a, b \in \mathbb{C}[G]$, one can observe that $f_{ab} = f_b * f_a$ with the definition above and the discrete measure. This makes the identification $\text{Fun}(G \to \mathbb{C}) \cong \mathbb{C}[G]^{opp}$ an isomorphism of algebras. We have identified L_g with the action of G on $\mathbb{C}[G]$ by left multiplication, so $L_g f_{ab} = f_{gab} = f_b * f_{ga} = f_b * L_g f_a$. The action of G by left-multiplication commutes with the action of $\mathbb{C}[G]^{opp}$ by right-multiplication, and the above lemma shows that the analogy holds when G is infinite.

Stone-Weierstrass

A critical ingredient in leveraging the algebraic structure of $C_{alg}(G)$ to prove its density in $L^2(G)$ is the Stone-Weierstrass theorem. Remarkably, the proof of this theorem can be boiled down to constructing a power series which converges uniformly to the absolute value. This establishes the connection between the density of other familiar function algebras, such as polynomials and trigonometric polynomials on a compact interval, and $C_{\text{alg}}(G)$. We also note that the compactness of G is an important ingredient here.

Theorem 11 (Dini's Theorem)**.** Let $\{f_n\}$ be a sequence of real, nonnegative functions in $C(I)$, where I is a compact interval, converging pointwise to a continuous function f and $f_n \leq f$. If $\{f_n(x)\}$ is monotone increasing for all $x \in I$, then $\{f_n\}$ converges uniformly to f on I.

Proof. Put $g_n = f - f_n$. For $\epsilon > 0$, consider the open subset $E_n = \{x \in I : g_n(x) < \epsilon\}$. Since $g_{n+1} \leq g_n$, $E_n \subseteq E_{n+1}$. Since $g_n(x) \to 0$ for all $x \in I$, $\{E_n\}$ forms an open cover of I. We choose a finite subcover, and since the sequence of sets is monotone increasing, the join of this subcover is E_N for some N. Thus for each $\epsilon > 0$ there exists N such that for $n \geq N$, $f(x) - f_n(x) < \epsilon$ for all $x \in I$, so $f_n \to f$ uniformly. □

Remark 18. One might look at the statement of Dini's theorem and wonder how it relates to the Monotone Convergence Theorem. In fact, the Monotone Convergence Theorem for measurable functions on a compact interval will follow from Dini's theorem by approximation with continuous functions via Lusin's Theorem.However, MCT holds more generally in a non-compact space, whereas Dini's theorem does not.

Lemma 10. Put $p_0(x) = 0$ and $p_{n+1}(x) = p_n(x) + \frac{1}{2}(x^2 - p_n(x)^2)$. Then $p_n(x) \to |x|$ uniformly on $[-1, 1]$. *Note that this appears as exercise 23 of chapter 7 in Baby Rudin [7].*

Proof. Let $d_n(x) = |x| - p_n(x)$. We will verify inductively that $d_n(x) = x^2 - p_n(x) \leq |x|\Big(1 - \frac{|x|}{2}\Big)^n$. We find that

$$d_{n+1}(x) = |x| - p_{n+1}(x) = d_n(x) - \frac{x^2 - p_n(x)^2}{2} \tag{51}$$

$$= d_n(x) - \frac{1}{2}d_n(x)(|x| + p_n(x)) \tag{52}$$

$$= d_n(x)(1 - \frac{|x| + p_n(x)}{2}) \tag{53}$$

$$= d_n(x)(1 - |x| + \frac{1}{2}d_n(x)) \tag{54}$$

Now we claim that $d_n(x) \leq |x|$. For the base case, $p_0 = 0$ and $p_1 = \frac{x^2}{2} \leq |x|$. Applying the inductive assumption, $d_{n+1}(x) \geq d_n(x)(1 - |x|) \geq 0$. This shows that d_n is monotone decreasing and positive, which proves the claim. Thus $d_n \leq |x|$, so $d_{n+1}(x) \leq d_n(x)(1 - \frac{|x|}{2}) = |x|(1 - \frac{|x|}{2})^{n+1}$. This completes the induction, showing that $p_n(x) \to |x|$ pointwise for $x \in [-1, 1]$. Applying Dini's theorem, the convergence is uniform. □

Corollary 12. Let K be compact and $A \subseteq C(K, \mathbb{R})$ be a unital subalgebra that is closed in the uniform norm. For any $\{f_1, \ldots, f_n\} \subseteq A$, $\max_i f_i \in A$ and $\min_i f_i \in A$.

Proof. First we prove the claim for two functions f, g. If $f = g$ there is nothing to do, so let $f \neq g$. We can write $\max\{f, g\} = \frac{f+g+|f-g|}{2}$ and $\min\{f, g\} = \frac{f+g-|f-g|}{2}$. Since K is compact, $a = \max_{x \in K} |f(x) - g(x)|$ exists, and thus we can take $f' = \frac{f}{a}$ and $g' = \frac{g}{a}$. By Lem. 10, since $h \equiv |f' - g'| \leq 1$, we can write h as a convergent power series, and since A is closed, we have $h \in A$. Therefore $\max\{f, g\} = \frac{f+g+ah}{2} \in A$. Similarly, $\min\{f, g\} = \frac{f+g-ah}{2} \in A$. We can extend this argument inductively to $\{f_1, f_2, \ldots, f_n\}$. This proves the claim. □

Theorem 13 (Stone Weierstrass)**.** Let K be a compact space, and let $A \subseteq C(K, \mathbb{R})$ be a unital subalgebra. Then if A separates points, it is dense in $C(K, \mathbb{R})$ with respect to uniform convergence.

Proof. Fix $\epsilon > 0$ and let $f \in C(K, \mathbb{R})$ be arbitrary. Then for each $s, t \in K$ there exists h such that $h(s) \neq h(t)$, let

$$h_{s,t}(v) = f(s) + (f(t) - f(s))\frac{h(v) - h(s)}{h(s) - h(t)}$$

Fix s and let $g_t = h_{s,t} - f$. Note that $g_t \in C(K, \mathbb{R})$. Put $U_t = g_t^{-1}((-\infty, \epsilon))$, and observe that U_t is open. Thus $K \subseteq \bigcup_{t \in K} U_t$, and since K is compact, choose $\{U_{t_i}\}_{i=0}^N$ to be a finite subcover. Set $h_s = \max_{t_i}\{h_{s,t_i}\}$. Next, take $g_s = f - h_s$ and note that g_s is continuous. Let $V_s = g_s^{-1}((-\infty, \epsilon))$, and observe that V_s is open. In the same way, we construct a cover $K \subseteq \bigcup_{s \in K} V_s$, and choose a finite subcover $\{V_{s_i}\}_{i=1}^M$. Now we have $u = \min_{s_i} \max_{t_i} h_{s_i,t_i}$ with $h_{s_i,t_i} \in A$, and clearly $f - \epsilon \leq u \leq f + \epsilon$. Using the lemma, $u \in \overline{A}$. This shows that $\overline{A} = C(K, \mathbb{R})$, which shows that A is dense in $C(K, \mathbb{R})$. □

Remark 19. If K is Hausdorff, then the converse holds as well.

Remark 20. Since polynomials form a subalgebra of continuous functions and $p(x) = x$ clearly separate points, the Weierstrass Approximation Theorem follows.

Corollary 14. Let K be compact and $A \subseteq C(K, \mathbb{C})$ be a unital, self-adjoint subalgebra. Then the previous theorem holds.

Proof. Let f in $C(K, \mathbb{C})$ be arbitrary. Notice that $A_{\mathbb{R}}$, the set of real-valued functions in A, is a subalgebra of A that is unital and separates points, and $\mathrm{Re}(f), \mathrm{Im}(f) \in A_{\mathbb{R}}$. Since A is self-adjoint, $f = \mathrm{Re}(f) + i\,\mathrm{Im}(f) \in \overline{A}$. This shows that $\overline{A} = C(K, \mathbb{C})$, so A is dense in $C(K, \mathbb{C})$. □

Corollary 15. By Lusin's Theorem, $C(G, \mathbb{C})$ is dense in $L^2(G)$, therefore A is dense in $L^2(G)$.

Peter-Weyl

Density of matrix coefficients

Theorem 16 (Peter-Weyl (part 1))**.** Let G be a compact, Hausdorff topological group. For each irreducible finite-dimensional representation V of G, fix an orthonormal basis $\{\psi_j^V\}_{j=1}^{\dim V}$ with respect to a $G-$invariant inner product. Then $\{\sqrt{\dim V}\rho_{ij}^V\}_{i,j,V\in\widehat{G}}^{\dim V}$ forms an orthonormal basis for $L^2(G)$.

Proof. Since $C_{\text{alg}}(G)$ is a unital subalgebra of $C(G,\mathbb{C})$, it is sufficient to show that it separates points. Thus we must show that there exists some finite-dimensional invariant subspace V such that $\rho^V(g)\neq\rho^V(e)$. For the sake of contradiction, suppose this is not the case. We have shown that T_ϕ is a compact operator, so by the spectral theorem, the $\lambda-$eigenspace V_λ of T_ϕ is such a subspace. The assumption therefore implies that G acts by the identity for every eigenspace V_λ of T_ϕ for any $\phi\in L^2(G)$ satisfying $\phi(g)=\overline{\phi(g^{-1})}$. Since $V_0\oplus\bigoplus_\lambda V_\lambda$ is dense in $L^2(G)$, this implies that $(L_g-1)f\in\ker T_\phi$, so $T_\phi(L_g-1)f=0$. It follows that $L_g(T_\phi f)=T_\phi f$.

However, let $U_e, U_{g^{-1}}$ be disjoint open neighborhoods of e, g^{-1}, using the fact that G is Hausdorff. Then by taking $U_e\cap U_e^{-1}$, since inversion is homeomorphic, we may consider U_e and $U_{g^{-1}}$ to be symmetric. Next put $U=U_e\cap U_{g^{-1}}g$, and notice that $U\cap Ug^{-1}=U_e\cap U_{g^{-1}}g\cap U_eg^{-1}\cap U_{g^{-1}}\subseteq U_e\cap U_{g^{-1}}=\emptyset$. Consider $f=\phi=\mathbb{1}_U$. Then

$$T_\phi f(x)=\int_G \mathbb{1}_U(h^{-1}x)\mathbb{1}_U(h)dh=\mu(h\in U: hx\in U)$$

Therefore $T_\phi f(e)=\mu(U)$, but $[L_gT_\phi f](e)=0$. Since $\{gU: g\in G\}$ is an open cover of G, we choose a finite subcover $\{g_iU\}_{i=1}^N$. By translation-invariance, $\mu(g_iU)=\mu(U)$, so $\mu(G)\leq N\mu(U)$. This shows that $\mu(U)>0$, which is a contradiction to the assumption, showing that the matrix elements separate points. By applying the Stone-Weierstrass theorem, we find that the matrix coefficients are dense in $C(G,\mathbb{C})$. Since $C(G,\mathbb{C})$ is dense in $L^2(G,\mathbb{C})$ by Lusin's Theorem, this also shows density in $L^2(G)$. As we have shown above, $C_{\text{alg}}(G)$ is equal to the linear span of the matrix coefficients of irreducible representations. This makes $\{\sqrt{\dim V}\rho_{ij}^V\}$ an orthonormal basis for $L^2(G)$. ☐

Corollary 17. If $L^2(G)$ is separable, then G has countably many non-isomorphic irreducible representations.

Proof. Let $S=\{e_i\}_{i\in I}$ be any orthonormal basis for $L^2(G)$. Since $L^2(G)$ is a separable metric space, S is separable. We see that if $i\neq j$, $\|e_i-e_j\|^2=\|e_i\|^2+\|e_j\|^2=2$. This shows that the only dense subset of S is S itself, so S is countable. ☐

Decomposition of the left-regular representation

Lemma 11. If (ρ,V) is a finite-dimensional representation of G, then for any $w\in V$, define $L^2(G)_{V,w}\equiv\operatorname{span}\{\rho_{v^*,w}^V : v\in V\}$ and $\psi_w: V\to L^2(G)_{V,w}$, via $\psi_w(v)=\rho_{v^*,w}^V$. Then ψ is an isomorphism of representations with the action of G on $L^2(G)_{V,w}$ by left-translation.

Proof. Take $\langle -,-\rangle$ to be a $G-$invariant inner product on V. If $h\in G$ is arbitrary, then

$$[L_g\psi_w(v)](h)=\langle\rho(g^{-1}h)w,v\rangle=\langle\rho(h)w,\rho(g)v\rangle=[\psi_w(\rho(g)v)](h)$$

This shows that ψ is a homomorphism of representations. ψ is clearly surjective, and comparing dimensions shows that ψ is invertible, so it is an isomorphism of representations. ☐

Corollary 18 (Peter-Weyl (part 2))**.** Under the action of G by left-translation, we have $L^2(G)\cong\widehat{\bigoplus}_{V\in\widehat{G}}V^{\oplus\dim V}$.

Proof. Let $V \in \widehat{G}$. We have shown that $(\rho^V, V) \cong (L, L^2(G)_{V,w})$. Let $\{w_i\}$ be an orthonormal basis for V with respect to the $G-$invariant inner product. By Schur orthogonality, we can form the direct sum $L^2(G)_V \equiv \bigoplus_{i=1}^{\dim V} L^2(G)_{V,w_i}$, and $L^2(G)_V \cong V^{\oplus \dim V}$ isometrically with appropriately rescaled inner product. We see that $\bigoplus_{V \in \widehat{G}} L^2(G)_V = C_{\text{alg}}(G)$ so by the Peter-Weyl theorem, it is dense in $L^2(G)$. Therefore taking the completion, $L^2(G) \cong \widehat{\bigoplus}_{V \in \widehat{G}} L^2(G)_V \cong \widehat{\bigoplus}_{V \in \widehat{G}} V^{\oplus \dim V}$. □

Representations in Hilbert space

Definition 25 (Integration in Hilbert spaces)**.** Suppose X is a probability measure space, H is a Hilbert space, and $f : X \to H$ is bounded, i.e. $\|f(x)\| \leq C$ for some C and all $x \in X$. Further suppose that $f_v(-) = \langle v, f(-)\rangle$ is measurable for all $v \in H$. Define the operator $T(-) = \int_X \langle -, f(x)\rangle d\mu(x)$. Take v to have unit norm. Then $|T(v)| \leq \int_X |\langle v, f(x)\rangle| d\mu(x) \leq \int_X \|f(x)\| d\mu(x) \leq C$. This shows that $\|T\|_{op} \leq C$. Thus $T : H \to \mathbb{C}$ defines a bounded linear functional, so there exists $F \in H$ such that $T(-) = \langle -, F\rangle$ by the Frechet-Riesz representation theorem. f is said to be scalarly integrable and we call F the Pettis integral of f.

Lemma 12 (Properties of the Pettis integral)**.** Let $f, g : X \to H$ be scalarly integrable functions and let α, β be complex scalars. Let T be a bounded linear operator on H, and let B be an orthonormal basis for H. Then for $\alpha, \beta \in \mathbb{C}$, the following properties hold:

(1) $\int_X \alpha f(x) + \beta g(x) d\mu(x) = \alpha \int_X f(x) d\mu(x) + \beta \int_x g(x) d\mu(x)$

(2) Tf is integrable and $T \int_X f(x) d\mu(x) = \int_X Tf(x) d\mu(x)$

(3) If F exists, then $\|F\| \leq \int_X \|f(x)\| d\mu(x)$.

(4) $u-$substitution

Proof. Let F, G be the integrals of f and g. Then for any $v \in H$,

$$\begin{aligned} \langle v, \alpha F + \beta G\rangle &= \overline{\alpha}\langle v, F\rangle + \overline{\beta}\langle v, G\rangle && (55)\\ &= \overline{\alpha} \int_X \langle v, f(x)\rangle d\mu(x) + \overline{\beta} \int_X \langle v, g(x)\rangle && (56)\\ &= \int_X \langle v, \alpha f(x) + \beta g(x)\rangle d\mu(x) && (57)\end{aligned}$$

This shows that $\alpha F + \beta G$ is the integral of $\alpha f + \beta g$. Next, let T be a bounded operator. Then $\|Tf(x)\| < \|T_{op}\|\|f(x)\|$, so Tf is scalarly integrable. Similarly, if T is bounded then it has an adjoint, so

$$\langle v, TF\rangle = \langle T^\dagger v, F\rangle = \int_X \langle T^\dagger v, f(x)\rangle d\mu(x) = \int_X \langle v, Tf(x)\rangle d\mu(x) \tag{58}$$

This shows that TF is the integral of Tf. Next, notice that

$$\|F\|^2 = |\langle F, F\rangle| = |\int_X \langle F, f(x)\rangle d\mu(x)| \leq \|F\| \int_X \|f(x)\| d\mu(x)$$

Lastly, u-substitution follows immediately from the standard integral. □

Corollary 19 (Peter-Weyl (Part 3))**.** Let ρ be a representation of G on a Hilbert space V. Then V^{fin} is dense in V.

Remark 21. This statement says that an analogue of complete reducibility holds for infinite-dimensional representations of compact topological groups.

Remark 22. The convolution operator makes $L^2(X)$ into an algebra, but this algebra has no unit. The would-be unit is the Dirac delta at the identity. This is not a proper function, but can be approximated arbitrarily closely. In particular, the density of the matrix coefficients allows us to approximate the identity by "smearing" the action of G with a representative function.

Proof. Suppose $v \in V$ is arbitrary. The map $f_v : g \mapsto \rho(g)v$ is norm-continuous by assumption. Since V is $G-$invariant, $f_v : G \to V$. Since G is compact, we have $M = \sup_{h\in G} \|f_v(h)\|_V < \infty$. Let $\phi \in C(G)^{\text{fin}}$ be arbitrary. Since ϕ is continuous and therefore bounded because G is compact, $h \mapsto \phi(h)f_v(h)$ satisfies the requirement for scalar integrability. Let $S_\phi : V \to V$ via

$$S_\phi v = \int_G \phi(h)\rho(h)v d\mu(h)$$

Since $\rho(g)$ is continuous, through $u-$substitution, we observe that

$$\rho(g)[S_\phi v] = \int_G \phi(h)\rho(gh)vdh = \int_G \phi(g^{-1}h)\rho(h)vdh = S_{L_g\phi}v \tag{59}$$

Since $(\phi, v) \to S_\phi v$ is bilinear and ϕ generates a finite-dimensional subspace, this shows that $S_\phi v \in V^{\text{fin}}$. Furthermore, since $\rho(e)v = v$, we may choose a neighborhood U of the identity with nonzero measure such that $\|\rho(g)v - v\|_V < \epsilon$ for all $g \in U$. As we have shown, $C(G)^{\text{fin}}$ is dense in $L^2(G)$, so we may find $\phi \in C(G)^{\text{fin}}$ such that $\|\phi - \delta_U\|_{L^2} < \epsilon/M$. We note that $\|\phi - \delta_U\|_{L^1} \leq \|\phi - \delta_U\|_{L^2}$. From the triangle inequality,

$$\|S_\phi v - f\|_V \leq \|S_\phi v - S_{\delta_U} v\|_V + \|S_{\delta_U} v - v\|_V \tag{60}$$

We can see that

$$\begin{aligned}
\|S_\phi v - S_{\delta_U} v\|_V &= \left\| \int_G (\phi(g) - \delta_U(g)) f_v(g) d\mu(g) \right\|_V && (61)\\
&\leq \int_G |\phi(g) - \delta_U(g)| \|f_v(g)\|_V d\mu(g) && (62)\\
&\leq M\|\phi(g) - \delta_U\|_{L^1} && (63)\\
&\leq \epsilon && (64)
\end{aligned}$$

Lastly, we can show that $S_{\delta_U} v$ approximates v:

$$\|S_{\delta_U} v - v\|_V = \frac{1}{\mu(U)} \left\| \int_U \rho(g)v - vd\mu(g) \right\|_V \leq \frac{1}{\mu(U)} \int_U \|\rho(g)v - v\|_V \, d\mu(g) \leq \epsilon$$

Thus $\|S_\phi v - v\|_V \leq 2\epsilon$, so v may be approximated by elements of V^{fin}. This shows that V^{fin} is dense in V. ☐

Corollary 20. All irreducible Hilbert space representations of G are finite-dimensional.

Proof. Let V be an irreducible infinite-dimensional Hilbert space representation of G. Then $V^{\text{fin}} = \{0\}$. This contradicts the fact that V^{fin} is dense in V. ☐

Corollary 21. If (ρ, V) is a unitary representation of G, then $V = \widehat{\bigoplus}_{U\in\widehat{G}} V_U$, where V_U is a direct sum of all the isomorphic copies of U occurring in V.

Proof. By Zorn's lemma, choose a maximal collection of orthogonal finite-dimensional irreducible representations, and organizing by isomorphism class, construct $W = \bigoplus_{U\in\widehat{G}} V_U$. Take $\overline{W}^\perp$. Since the representation is unitary, $\overline{W}^\perp$ is orthogonal to V^{fin}. This contradicts the density of V^{fin}. ☐

Applications

Linearity of compact Lie groups

The Peter-Weyl theorem remarkably holds that every compact Lie group is a linear group, i.e. a subgroup of $GL(n, \mathbb{C})$ for some finite n. This is equivalent to the statement that G has a faithful (i.e. injective) finite-dimensional representation $\rho : G \to \mathrm{End}(\mathbb{C}^n)$, from which the density of matrix coefficients follows immediately from the Stone-Weierstrass Theorem. As a result, this statement it often called the Peter-Weyl theorem in the context of Lie groups.

Theorem 22 (Closed Subgroup Theorem)**.** Let G be a Lie group. $H \leq G$ is a topologically closed subgroup of G iff H is an embedded Lie subgroup of G. See page 28 of Brocker and Dieck [1].

Lemma 13. Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. Then the exponential map $\exp : \mathfrak{g} \to G$ via $v \mapsto \gamma_v(1)$, where γ_v is the integral curve of v passing through e, is a local diffeomorphism [2].

Lemma 14. Let G be a Lie group, and let G^0 be the connected component of G containing the identity. Then G^0 is a normal subgroup of G, and G/G^0 is discrete. Furthermore, $|G/G^0|$ equals the number of connected components of G. The proof can be found in chapter 2 of Kirillov [5].

Lemma 15 (Descending Chain Property)**.** Let G be a compact Lie group of finite dimension. Let $G > K_1 > K_2 > \dots$ be a chain of topologically closed Lie subgroups of G. Then then chain terminates at $K_N = \{e\}$ for some finite N.

Proof. Suppose that $\dim K_n = \dim K_{n+1}$ for some n. By the closed subgroup theorem, the inclusion $K_{n+1} \hookrightarrow K_n$ is an embedding. From this, $\mathfrak{k}_{n+1}$, the Lie algebra of k_{n+1}, is isomorphic to $\mathfrak{k}_n$, so by the assumption $\mathfrak{k}_{n+1} = \mathfrak{k}_n$. The exponential map $\exp : \mathfrak{k}_{n+1} \to K_{n+1}$ is a local diffeomorphism, so K_{n+1}^0 is an open subgroup of K_n^0. By the closed subgroup theorem, it is also closed, so $K_{n+1}^0 = K_n^0$ by connectedness. Therefore $K_{n+1}/K_{n+1}^0 \leq K_n/K_n^0$. Since $K_n, K_{n+1} < G$ are compact and thus have finitely many connected components, both are finite groups, so by Lagrange's theorem, $|K_{n+1}/K_{n+1}^0|$ divides $|K_n/K_n^0|$. This shows that either K_{n+1} has less connected components than K_{n+1}^0 or $\dim K_{n+1} < \dim K_n$. Therefore, the chain terminates with some K_N with $\dim K_N = 0$ and K_N has one connected component, which implies that $K_N = \{e\}$. □

Theorem 23 (Peter-Weyl for compact Lie groups)**.** Every compact Lie group is isomorphic to a subgroup of $\mathrm{GL}(n, \mathbb{C})$ for some finite n.

Proof. We have already shown that the matrix coefficients of G separate points. Take (ρ_1, V_1) to be some non-trivial representation of G with $g \notin \ker \rho_1$. Then $\ker \rho_1 \leq G$ is a closed subgroup of G, and by the Closed Subgroup Theorem, it is a Lie subgroup of G. Then we may choose $h \in \ker \rho_1$ and some representation (ρ_2, V_2) of G such that $h \notin \ker \rho_2$. We see that $\ker \rho_1 \geq \ker \rho_1 \oplus \rho_2$, and the containment is proper. Proceeding iteratively, we construct a chain which terminates with $\ker \rho_1 \geq \ker \rho_1 \oplus \rho_2 \geq \dots \geq \ker \bigoplus_{n=1}^N \rho_n = \{e\}$. Since the kernel of $\bigoplus_{n=1}^N \rho_n$ is trivial, this representation is faithful, and thus it is a group isomorphism to some subgroup of $\mathrm{GL}(\sum_{n=1}^N \dim V_n, \mathbb{C})$. □

Remark 23. This theorem is sometimes referred to as the Peter-Weyl theorem, because the Stone-Weierstrass theorem directly implies the density of the matrix coefficients.

Fourier analysis

Definition 26. We identify $\mathrm{End}(V) \cong V \otimes V^*$ by $(v \otimes \omega^*)(u) = \omega^*(u)v$. The action of $G \times G$ on $\mathrm{End}(V)$ can be defined as the action on $V \otimes V^*$ using the definition of the contragredient and tensor product representation as defined previously:

$$\rho^{V \otimes V^*}((g, h))(v \otimes w^*) = \rho^V(g)v \otimes (\rho^V(h)w)^*$$

We can see that $V \otimes V^*$ is an irreducible $G \times G$-representation if V is an irreducible $G-$representation. Furthermore, let $u \in V$ be arbitrary. We see that

$$[\rho^V(g)v \otimes (\rho^V(h)w)^*]u = \rho^V(g)v\langle u, \rho^V(h)w\rangle \tag{65}$$
$$= \rho^V(g)v\langle \rho^V(h^{-1})u, w\rangle \tag{66}$$
$$= (\rho^V(g)v \otimes w^*)\rho^V(h^{-1})u \tag{67}$$

Extending by linearity, if $A \in \mathrm{End}(V)$ is arbitrary, then $G \times G$ acts on A via

$$\rho^{\mathrm{End}(V)}((g,h))A = \rho^V(g)A\rho^V(h^{-1}) \tag{68}$$

Proposition 11. Let V represent G. The map $m : \mathrm{End}(V) \to L^2(G)_V$ via $m(v\otimes\omega^*) = \sqrt{\dim V}\rho_{\omega^*,v}$ is an isomorphism of representations of $G\times G$, with the action of $G\times G$ on $L^2(G)$ by right- and left-translation. Furthermore, this map is an isometry with the inner product $(A,B) \mapsto \langle A, B\rangle_{HS(V)} \equiv \mathrm{Tr}_V(B^\dagger A)$.

Proof. We may show these claims by direct computation. Let (ρ, V) be a representation of G and $\omega \in V^*$, $v \in V$, $(g,h) \in G \times G$ be arbitrary. Then

$$L_gR_hm(v \otimes \omega)(k) = \sqrt{\dim V}L_gR_h\omega(\rho^V(k)v) \tag{69}$$
$$= \sqrt{\dim V}\omega(\rho^V(g^{-1}kh)v) \tag{70}$$
$$= \sqrt{\dim V}\omega(\rho^V(g^{-1})\rho^V(k)\rho^V(h)v) \tag{71}$$
$$= \sqrt{\dim V}(\rho^{V^*}(g)\omega)(\rho^V(h)v) \tag{72}$$
$$= m(\rho^V(h)v \otimes \rho^{V^*}(g)\omega)(k) \tag{73}$$
$$= m(\rho^{V\otimes V^*}((g,h))(v \otimes \omega))(k) \tag{74}$$

To check the isometry, it suffices to check basis elements. Fixing a orthonormal basis $v_1, \ldots, v_n$ with respect to the $G-$invariant inner product for V, we have

$$\langle\sqrt{\dim V}\rho^V_{ij}, \sqrt{\dim V}\rho^V_{kl}\rangle_{L^2} = \delta_{ik}\delta_{jl} = \mathrm{Tr}\{(\beta v_l \otimes v_k^*)^\dagger(\alpha v_j \otimes v_i^*)\}$$

$\square$

Corollary 24 (Peter-Weyl (part 4))**.** The map

$$m : \widehat{\bigoplus_{V\in\widehat{G}}} \mathrm{End}(V) \xrightarrow{\sim} L^2(G) \ ,$$

as defined previously, where the completion is taken with respect to the inner product defined above, is an isomorphism of representations of $G \times G$.

Proof. By Schur orthogonality of the matrix coefficients of non-isomorphic representations, the mapping is isometric. As such, it is injective. We have shown that $C_{\mathrm{alg}}(G) = \bigoplus_{V\in\widehat{G}} \mathrm{span}\{\rho^V_{ij}\}_{ij} = m(\bigoplus_{V\in\widehat{G}} V \otimes V^*)$. Since m is a continuous, continuously-invertible map, the image of the completion of the direct sum is a closed subspace of $L^2(G)$. By continuity, m is also isometric on the completion of the direct sum. By the projection theorem, we may decompose $L^2(G)$ into the image and its orthogonal complement. By the Peter-Weyl theorem, $C_{\mathrm{alg}}(G)$ is dense in $L^2(G)$, which requires that the orthogonal complement is trivial. This shows that m is also surjective, proving the claim. $\square$

Definition 27 (Fourier components)**.** Given $f \in L^2(G)$, we define the Fourier component of f in $V \in \widehat{G}$ as $\hat{f}(V) = \pi_{\mathrm{End}(V)}m^\dagger(f)$.

Corollary 25. The matrix representation of $\hat{f}(V)$ can be obtained explicitly via

$$[\hat{f}(V)]_{ij} = \langle f, \sqrt{\dim V}\rho^V_{ji}\rangle_{L^2}$$

Or equivalently,

$$\hat{f}(V) = \sqrt{\dim V}\int_G f(g)\rho^V(g^{-1})d\mu(g)$$

Proof. By the Peter-Weyl theorem, we have the $L^2(G)$ expansion $f = \sum_{V\in\widehat{G}}\sum_{ij}\alpha^V_{ij}\sqrt{\dim V}\rho^V_{ji}$. We see that $m^\dagger : f \mapsto \sum_{V\in\widehat{G}}\sum_{ij}\alpha^V_{ij}v_i\otimes v_j^*$, so

$$\hat{f}(V) = \pi_{\mathrm{End}(V)}m^\dagger(f) = \sum_{ij}\alpha^V_{ij}v_i\otimes v_j^* \tag{75}$$

This gives $[\hat{f}(V)]_{ij} = \alpha^V_{ij}$. By Fubini's Theorem and the orthogonality relations, we find that

$$[\hat{f}(V)]_{ij} = \alpha^V_{ij} = \langle f, \sqrt{\dim V}\rho^V_{ji}\rangle_{L^2}$$

as claimed. □

Corollary 26 (Fourier Inversion)**.** We can write $m(\hat{f}(V)) = \sqrt{\dim V}\langle\hat{f}(V),(\rho^V)^\dagger\rangle_{HS(V)}$, and we have the $L^2(G)$ expansion $f(g) = \sum_V\sqrt{\dim V}\langle\hat{f}(V),\rho^V(g)^\dagger\rangle_{HS(V)}$.

Proof. We have $\hat{f}(V) = \sum_{ij}[\hat{f}(V)]_{ij}v_i\otimes v_j^*$, so

$$\begin{aligned} m(\hat{f}(V))(g) &= \sum_{ij}[\hat{f}(V)]_{ij}\sqrt{\dim V}\rho^V_{ji}(g) \\ &= \sqrt{\dim V}\,\mathrm{Tr}\{\rho^V(g)\hat{f}(V)\} \\ &= \sqrt{\dim V}\langle\hat{f}(V),\rho^V(g)^\dagger\rangle_{HS(V)} \end{aligned}$$

Since $\pi_{\mathrm{End}(V)}$ form a complete set of mutually orthogonal projectors, we can write $Id = \sum_{V\in\widehat{G}}\pi_{\mathrm{End}(V)}$, and thus we obtain the Fourier inversion formula

$$f = m(\sum_{V\in\widehat{G}}\pi_{\mathrm{End}\,V}m^\dagger(f)) = \sum_{V\in\widehat{G}}m(\hat{f}(V)) = \sum_{V\in\widehat{G}}\sqrt{\dim V}\langle\hat{f}(V),(\rho^V)^\dagger\rangle_{HS(V)}$$

□

Corollary 27 (Plancherel identity)**.** From the unitarity of m, we have

$$\|f\|_{L^2} = \sum_{V\in\widehat{G}}\|\hat{f}(V)\|^2_{HS(V)}$$

Lemma 16. If $B\in\mathrm{End}(V)$, we have $\overline{m(B)(g)} = m(B^\dagger)(g^{-1})$.

Proof. Consider $m(B)(g) = \sqrt{\dim V}\sum_{ij}B_{ij}\rho^V_{ji}(g)$. Then we have

$$\overline{m(B)(g)} = \sum_{ij}\overline{B_{ij}}\overline{\rho^V_{ji}(g)} = \sum_{ij}B^\dagger_{ji}\rho^V_{ij}(g^{-1}) = m(B^\dagger)(g^{-1})$$

□

Proposition 12 (Convolution Identity)**.** If $L^2(G)$ is considered as an algebra with respect to the convolution $(f*g) = \int_G f(h)g(xh^{-1})d\mu(h)$, then m is also an algebra isomorphism.

Proof. It suffices to check for $A, B \in \mathrm{End}(V)$. Then

$$m(AB) = \langle AB, \overline{\rho^V} \rangle_{HS(V)}$$

$$\begin{aligned}
[m(A) * m(B)](g) &= \int_G m(A)(h) m(B)(gh^{-1}) dh \\
&= \int_G m(A)(h) \overline{R_{g^{-1}} m(B^\dagger)}(h) dh \\
&= \langle m(A), R_{g^{-1}} m(B^\dagger) \rangle_{L^2} \\
&= \langle m(A), m(\rho^V(g^{-1}) B^\dagger) \rangle_{L^2} \\
&= \langle A, \rho^V(g^{-1}) B^\dagger \rangle_V \\
&= \langle AB, \rho^V(g)^\dagger \rangle_V \\
&= m(AB)(g)
\end{aligned}$$

if we let $A = \hat{f}(V)$ and $B = \hat{g}(V)$ and apply $m^\dagger$ to both sides, we find

$$\widehat{f * g}(V) = \hat{f}(V) \hat{g}(V) \tag{76}$$

□

Special functions: Wigner functions and spherical harmonics

In quantum mechanics, the matrix coefficients of the irreps of SU(2) are called Wigner functions. Both the Schrodinger equation (in quantum mechanics) with a spherically symmetric potential and the Laplace equation (in electrodynamics) are separable differential equations involving the Laplace operator $\Delta = \partial_i \partial^i$ (where we use the summation convention going forward unless explicitly stated otherwise.) This operator commutes with rotation. In particular, a solution ψ to both differential equations can be separated into $\psi(\theta, \phi, r) = R(r)\Theta(\theta, \phi)$, so we can focus on decomposing $\Theta \in L^2(S^2)$ into irreducible representations of SO(3). Since $S^2 \cong \mathrm{SO}(3) \backslash \mathrm{SO}(2)$, this naturally relates $L^2(S^2)$ to the left-regular representation of SO(3). Here, we will attempt to give a short introduction to angular momentum and illustrate how the spherical harmonics, a common family of special functions, arise from the Peter-Weyl theorem.

Angular momentum and irreps of SO(3)

In quantum mechanics, quantum states are elements of a Hilbert space $\mathcal{H}$, and observable quantities are self-adjoint operators on $\mathcal{H}$. If $\psi \in \mathcal{H}$ is a quantum state and X is an observable, then the quantity $\psi^\dagger X \psi \equiv \langle X \rangle_\psi$ is called the expectation value of X, and this is the quantity that is observed in experiment. The self-adjointness of X guarantees that this quantity is real. If (X^1, X^2, X^3) is a vector of observables corresponding to position on the x, y, and z axes, to represent a rotation $R \in \mathrm{SO}(3)$ at the level of quantum states, we wish to find an operator $\mathcal{D}(R) \in \mathrm{SU}(2)$ such that when $\psi' = \mathcal{D}(R)\psi$, we have $\langle X^i \rangle_{\psi'} = R^i_j \langle X^j \rangle_\psi$. The operators X^1, X^2, X^3 are traceless, and mathematicians will be familiar with this as the double-cover of SO(3) by SU(2), where $\mathcal{D}(R) \in \mathrm{SU}(2)$ satisfies $\mathcal{D}^\dagger(R) X^i \mathcal{D}(R) = R^i_j X^j$.

This makes $\mathcal{H}$ into a representation of SU(2). The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ the vector space of traceless 2×2 complex matrices. Since $\mathfrak{sl}(2, \mathbb{C})$ is the complexification of $\mathfrak{su}(2)$, the irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$ are in direct correspondence with those of $\mathfrak{su}(2)$, and by extension, with those of SU(2). A common basis for $\mathfrak{sl}(2, \mathbb{C})$ is $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), f = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. These are known to physicists as the spin-raising operator, spin-lowering operator, and z-component of spin respectively. The Lie algebra is then defined by the brackets $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. Let ρ be a representation of $\mathfrak{sl}(2, \mathbb{C})$

in a finite-dimensional vector space V. If v is an eigenvector of $\rho(h)$ with eigenvalue λ, then from the commutators, $\rho(e)v$ is an eigenvector with eigenvalue $\lambda+2$ and $\rho(f)v$ is an eigenvector with eigenvalue $\lambda-2$. This is what motivates the terms spin-raising and spin-lowering. Since V is finite-dimensional, then this implies that the greatest eigenvector v with eigenvalue λ satisfies $\rho(e)v = 0$. Defining $v^k \equiv \frac{\rho^k(f)}{k!}v$, we must also have $v^{n+1} = 0$ for some n. By using the brackets, we work out $\rho(e)v^k = (\lambda - k + 1)v^{k-1}$. For $k = n+1$, we must then have $\lambda - n = 0$. Thus $\lambda = n$ is a non-negative integer, and this constructs all of the irreducible representations of SU(2).

Wigner functions and spherical harmonics

A common basis for the Lie algebra of SU(2) is $J_z = h$, $-iJ_x = e + f$, $-iJ_y = e - f$. Physicists refer to these as the z, x and y components of angular momentum respectively. We can parameterize a rotation via a unit vector $\hat{n}$ and an angle ϕ by exponentiation:

$$\mathcal{D}(\hat{n}, \phi) \equiv \exp\Big(-i\hat{n} \cdot \vec{J}\phi\Big) \tag{77}$$

where $\vec{J} = (J_x, J_y, J_z)$. Then we define the Wigner functions [8] to be $\mathcal{D}^l_{mm'}(\hat{n}, \phi) = (v^m)^\dagger \mathcal{D}(\hat{n}, \phi) v^{m'}$, where $v \in V[l]$ is the highest-weight vector as defined in the previous section. These functions are fundamental to atomic physics, as they describe the way quantum states rotate. We can see that these are explicitly the matrix coefficients appearing in the Peter-Weyl theorem.

For a more practical application and familiar example, we turn toward the spherical harmonics. Going back to our definition of the double-cover $\mathrm{SU}(2) \to \mathrm{SO}(3)$, we can now interpret $\exp(i\pi\rho(h))$ as a rotation by 2π about the z axis, and the single-valuedness of this rotation implies that $V[n]$ may be lifted to a representation of SO(3) iff n is an even integer. We return to the representation of SO(3) in $L^2(S^2)$ by rotations. For simplicity, let $G = \mathrm{SO}(3)$ and $H = \mathrm{SO}(2)$. From the Peter-Weyl theorem, we have $L^2(G) \cong \bigoplus_{k=1}^\infty V[2k] \otimes V^*[2k]$. If we consider the subspace $L^2(G/H) \leq L^2(G)$ of functions which are constant on left-cosets, then these functions are invariant under the right-action of H. This shows that $L^2(G/H) \cong \bigoplus V[2k] \otimes W^*[2k]$, where $W^*[2k]$ is the SO(2)-invariant subspace of $V[2k]$. If h spans $\mathfrak{so}(2) \leq \mathfrak{so}(3)$, then we similarly find that $V[2k]$ has exactly one eigenvector of h with eigenvalue zero, so $W[2k]$ contains a single copy of the trivial representation. This shows that

$$L^2(S^2) \cong \bigoplus_k V[2k] \otimes \mathrm{triv} \cong \bigoplus_k V[2k] \tag{78}$$

where $V[2k]$ are invariant subspaces of functions on the sphere. The spherical harmonics are defined as

$$Y^l_m = \sqrt{\frac{2l+1}{4\pi}} v^m \tag{79}$$

where v is the highest-weight vector of $V[2l]$ as defined in the previous section (the normalization is by convention)[3]. We can then relate the spherical harmonics to the Wigner functions using the same method we employed with the matrix coefficients. WLOG take h to represent a rotation about the z axis. Then by left-translation, $Y^l_m(\hat{n}) = R^\dagger(\hat{n}) Y^l_m(\hat{z})$, where $R(\hat{n})$ is a rotation that takes $\hat{z}$ to $\hat{n}$. By definition, $R^\dagger(\hat{n}) Y^l_m(\hat{z}) = (D^l_{mm'}(R))^* Y^l_{m'}(\hat{z})$. Since h represents a rotation about $\hat{z}$, we have $h = \frac{\partial}{\partial\phi}$, where ϕ is the angle coordinate about $\hat{z}$. Then Y^l_m must satisfy the equation $\frac{\partial}{\partial\phi} Y^l_m = l(l-m) Y^l_m$, which sets $Y^l_m \propto e^{(n-2m)h\phi}$. For well-definedness, this shows that Y^l_m must vanish at $\theta = 0$ unless $m = l$. This sets $Y^m_l(\hat{n}) = (D^l_{ml}(R))^* Y^l_l(\hat{z})$. Since the latter is just a constant, this shows that Y^l_m is proportional to D^l_{ml}, showing how the spherical harmonics are related to the Wigner functions.

[3]This definition differs from the traditional one by a small detail; usually v^m is taken to be the eigenvector of h with eigenvalue $2m$, but here v^m has eigenvalue $2l - 2m$.

References

[1] Theodor Bröcker and Tammo Tom Dieck. *Representations of compact Lie groups*, volume 98. Springer Science & Business Media, 2013.

[2] Daniel Bump et al. *Lie groups*, volume 225. Springer, 2004.

[3] Roger W Carter, Ian G MacDonald, Graeme B Segal, and M Taylor. Lectures on lie groups and lie algebras. 1995.

[4] Jorge Blanco Herrera. The haar measure existence and uniqueness $\pm\epsilon$. 2024.

[5] Alexander A Kirillov. *An introduction to Lie groups and Lie algebras*, volume 113. Cambridge University Press, 2008.

[6] Walter Rudin. *Real and complex analysis, 3rd ed.* McGraw-Hill, Inc., USA, 1987.

[7] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1964.

[8] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics.* Cambridge University Press, 2020.

[9] Barry Simon. *Representations of finite and compact groups.* Number 10. American Mathematical Soc., 1996.

[10] Boris Weisfeiler. Abstract homomorphisms of big subgroups of algebraic groups. In *Topics in the Theory of Algebraic Groups*, volume 10, pages 135–182. University of Notre Dame, 1982.