M 392C Representation Theory, Fall 2025: Dr. David Ben-Zvi's lectures, TEXed by Sai Sivakumar. The main goal of these notes was to clarify and even expand on several of the points David made during the lectures. As a warning, I made minimal effort to reorder the content of the course, and many digressions in these notes have additional details not discussed in class.

Contents

Lecture 1 August 26

Definitions

Let G be a group and k be a field. A representation of G is a vector space $V \in \mathbf{Vect}_k$ with a linear action of G on V. (For convenience, let's assume representations are finite-dimensional vector spaces unless otherwise stated.)

This is to say there is a homomorphism $G \xrightarrow{\rho} \mathrm{Aut}(V)$ so that we may define the action of G on V to be $g \cdot v := \rho(g)v$ or alternatively a product $G \times V \to V$ given by $(g, v) \mapsto \rho(g)v$ that is linear in V and associative in G, among other properties. (We often suppress the $\cdot$ in the action/product.)

Note $\mathrm{Aut}(V) =: \mathrm{GL}(V) \cong \mathrm{GL}_{\dim V}(k)$, so we could even think of the elements of G as matrices if we fix a basis of V.

The (finite-dimensional) representations (V, ρ) of G over a field k form the objects of a category $\mathbf{Rep}_k(G)$ whose morphisms are linear transformations which commute with the action of G; i.e., $\mathrm{Hom}_G(V, W)$ consists of the linear transformations $T \colon V \to W$ for which $g(Tv) = T(gv)$ for all $g \in G$ and $v \in V$. (These are also called intertwining operators, G-linear maps, or G-equivariant maps, etc. Also note that $\mathrm{Hom}_G(V, W)$ is merely a vector space, but see this MSE post.)

The field k does matter, but for the majority of what follows k will be $\mathbb{C}$. Sometimes it may be possible to replace $\mathbb{C}$ with any algebraically closed field, but the characteristic of the field is rather important (see modular representation theory for example).

Variants of representations

We can add adjectives in various places to the definitions above to get different kinds of representations.

For example, if V has an inner product we can ask about linear actions of G on V that preserve that inner product. If V is a Hilbert space we can ask about linear actions of G on V for which $\rho(g)$ is a unitary operator; these are the unitary representations of G.

If G has a topology and V is a topological vector space, we say (V, ρ) is a continuous representation if the product map $G \times V \to V$ defined before is continuous. If G has a smooth structure (e.g. is a Lie group) or is an algebraic group, we can ask about smooth or algebraic representations, respectively.

Aspects of representations we study

1. Irreducible representations (henceforth called "irreps"): An irrep of G are the representations V which have no invariant (or stable some might say) subspaces under the group action; that is, if W is a G-invariant subspace of V then W is either 0 or V. These are like the "atoms" in representation theory.

   For a given group G one goal of representation theory is to classify up to isomorphism its irreps. One approach is to try to attach numerical invariants to representations and see if they help to classify irreps, and generalizations of this idea lead to character theory. Sometimes there are "isotopes" that are not isomorphic but nevertheless cannot be distinguished by certain invariants.

   Another thing we do is to take a subgroup H of G and study how irreps of G decompose when the group action is restricted to H. On the other hand, we can also look at how to build bigger representations or "molecules" out of the atomic irreps via extension. Of course, it is easy to take direct sums of irreps but depending on the context it may be possible to obtain indecomposable reps which are extensions of irreps but do not decompose into a direct sum of irreps (so to summarize, irreducible implies indecomposable but not the other way around in general).

   Maschke's theorem implies that indecomposable representations of finite groups over fields with characteristic not dividing the order of the group are irreducible. Alternatively, the theorem implies that in this setting, all short exact sequences of representations split.

   For example, the indecomposable representations over $\mathbb{C}$ coincide with the irreps when G is finite. On the other hand, the shearing representation

   $$\begin{aligned} \mathbb{Z} &\to \mathrm{GL}_2(\mathbb{R}) \\ 1 &\mapsto \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \end{aligned}$$

   preserves the horizontal axis, so it is not an irrep, but is not decomposable since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is not diagonalizable. The representations (k^n, ρ) of $\mathbb{Z}$ are given by specifying an invertible $n \times n$ matrix, and any two such representations are isomorphic if the matrices specifying them are conjugate. If k is algebraically closed, the indecomposable representations in this setting correspond to Jordan blocks. Of course, every representation is the direct sum of indecomposables, and in this setting the Jordan normal form of an $n \times n$ matrix would describe the decomposition of k^n into indecomposables.

2. Harmonic analysis is the study of naturally appearing "large" representations; we would like to perform some kind of "spectroscopy" to determine what representations appear within them.

   If a group G acts on some object X (which could be a set, manifold, or otherwise a "geometric" object), we say X has some symmetries which we would like to "linearize". We achieve this by considering the k-valued functions on X for some field k; if X has additional structure we can restrict to functions on X which interact with that structure (e.g. measurable/continuous/etc maps)

   From a right action of G on X, function spaces on X inherit a natural linear left action of G; one is given by $gf(x) = f(xg)$. These representations are typically very decomposable or reducible. The usual birthplace of harmonic analysis is $G = X = \mathbb{R}^n$, as it acts on itself by translation. Much of the harmonic analysis on $\mathbb{R}^n$ can be discovered without using representation theory, but if we want to do a similar analysis on other spaces, representation theory is usually required.

3. Different groups give rise to different phenomena. Below are some examples of groups we will see again.

	compact	noncompact
Abelian	S^1	$\mathbb{R}$
non-Abelian	$\mathrm{SO}_3(\mathbb{R})$	$\mathrm{SL}_2(\mathbb{R})$, $\mathrm{SL}_2(\mathbb{C})$

It is a theorem that irreps of compact groups coincide with their indecomposables, and that irreps of Abelian groups are all one-dimensional.

The representation theory of S^1 leads to the theory of Fourier series, and in a similar way we can recover the Fourier transform from the representation theory of $\mathbb{R}$.

The Lie group $\mathrm{SO}_3(\mathbb{R})$ can be thought of as the group of rotations of a 2-sphere. One nice result is that the spherical harmonics are basis functions for the irreps of $\mathrm{SO}_3(\mathbb{R})$, which occur naturally as the atomic orbitals (see the Wikipedia article on spherical harmonics).

The representations of the groups $\mathrm{SL}_2(k)$ for various fields k appear all over math. We can study special functions like the Bessel or hypergeometric functions, or even modular forms. The hard Lefschetz theorem in algebraic geometry says that $\mathrm{SL}_2(\mathbb{C})$ acts on $H^*(X, \mathbb{C})$ for X a nice enough smooth projective variety. In physics, the special linear group sort of appears in the Lorentz group $\mathrm{SO}(1,3)^+ = \mathrm{PSL}_2(\mathbb{C})$.

Spectral theory

A slogan for what is to come: "commutativity implies geometry".

Let $k = \mathbb{C}$ and X a set. Then the complex-valued functions on X form a commutative algebra. This is some example of a functor suggestively called $\mathcal{O}$ from the category of some kind of geometric objects to commutative algebras.

To expand on the previous idea, here is a motto originating from Gelfand and Grothendieck's work:

1. Any commutative ring should be thought of as functions on some space.

   That is, there is some functor going from commutative rings to some category of geometric objects that realizes this idea. In particular we should be able to adjust the functor and its source/target to obtain an equivalence of categories.

$$\text{commutative rings} \longleftrightarrow \text{geometry}$$

2. Once we are in the situation where we have an equivalence of categories between commutative rings and geometric objects, we should further obtain a correspondence between the modules over a ring R and sheaves (special families of vector spaces or Abelian groups) on the corresponding geometric object X to R.

$$R\text{-modules} \longleftrightarrow \text{sheaves on } X$$

For example, if X is a finite set, the corresponding ring R is the finite-dimensional commutative algebra of complex-valued functions on X. This algebra is semisimple with $R = \bigoplus_i \mathbb{C}e_i$ where $e_i e_j = \delta_{ij}$. We can think of the e_i as delta/indicator functions on points of X.

An R-module M has the decomposition $M = \bigoplus_i e_i M$, which corresponds to a sheaf on X where at each point of X we imagine the corresponding module $e_i M$ lying on it:

$$\begin{array}{ccccc} M & e_1M & e_2M & \cdots & e_nM \\ X & \bullet & \bullet & \cdots & \bullet \end{array}$$

A short word about the group algebra

A representation V of a group G is given by a group homomorphism $G \xrightarrow{\rho} \mathrm{Aut}(V)$. Since $\mathrm{Aut}(V)$ is contained in $\mathrm{End}(V)$, a k-algebra, there is an object $k[G]$ called the group algebra for which the homomorphism ρ factors through $k[G]$; that is, the diagram

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, hook] & \mathrm{End}(V) \\ k[G] \arrow[ur, dashed, "\overline{\rho}"'] & \end{tikzcd}$$

commutes. One explicit description of $k[G]$ is the set of finite linear combinations of elements of G with coefficients in k:

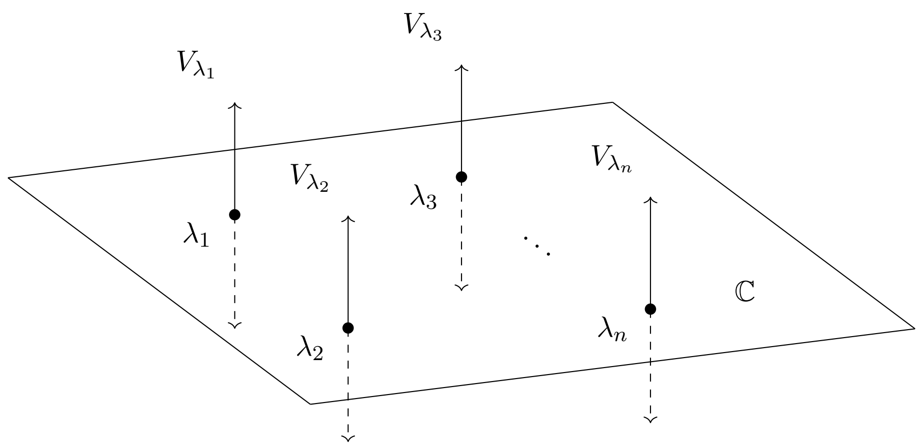
$$k[G] = \left\{ \sum_{g \in G} c_g g \;\middle|\; c_g \in k,\ \text{all but finitely many } c_k \text{ are zero} \right\}.$$

The addition and multiplication in $k[G]$ are defined using the addition in k and multiplication in G, respectively. We will return to the group algebra when investigating representations from a module-theoretic point of view.

Lecture 2 August 28

Spectral theory over $\mathbb{C}$

Consider an operator T on a finite-dimensional complex vector space V. The spectrum of T, $\sigma(T)$, is a finite subset of $\mathbb{C}$ consisting of the eigenvalues of T. To each $\lambda \in \sigma(T)$ there is a corresponding eigenspace V_λ of V, and $V = \bigoplus_{\lambda\in\sigma(T)} V_\lambda$. We can visualize placing each of the eigenspaces V_λ above each point λ in the spirit of sheaf theory:



On $\mathbb{C}$, the coordinate function x where $x(z) = z$ has the same action as T on the eigenspaces V_λ:

$$Tv = \sum_{\lambda\in\sigma(T)} Tv_\lambda = \sum_{\lambda\in\sigma(T)} \lambda v_\lambda = \sum_{\lambda\in\sigma(T)} x(\lambda)v_\lambda = x\cdot\left(\sum_\lambda v_\lambda\right) = x\cdot v$$

In this sense T corresponds to $x \in \mathrm{Fun}(\mathbb{C})|_{\sigma(T)}$ (complex-valued functions on $\mathbb{C}$, restricted to $\sigma(T)$.).

Now consider the commutative ring $R = k[T_1,\ldots,T_n]/(f_1,\ldots,f_m)$ and an R-module M. We would like to simultaneously diagonalize the action of the T_i on M, or rather, diagonalize the action of R on M, which amounts to finding a basis of M where R acts diagonally. Assume we can do this.

We should attach a set $X = \mathrm{Spec}(R)$ to R called the spectrum of R for which we can decompose M into a direct sum of modules M_{x_i}, each summand lying over their corresponding point $x_i \in X$:

$$\begin{array}{ccccccccc} M & = & M_{x_1} & \oplus & M_{x_2} & \oplus\cdots\oplus & M_{x_\ell} \\ X & & x_1 & & x_2 & \cdots & x_\ell \end{array}$$

Furthermore, we form an assignment $R \xrightarrow{\varphi} \mathrm{Fun}(X)$ for which we can recast the action of R on M through this assignment:

$$rm = \varphi(r)\sum_{x\in X} m_x = \sum_{x\in X}\varphi(r)|_{\{x\}}m_x$$

The idea here is to turn what was an algebraic notion of rings acting on modules to thinking about sheaves on a particular space which "diagonalize" the ring action.

Some examples of the algebra-geometry dictionary

An important philosophy we've been looking at the broad strokes of is this dictionary between algebra and geometry, specifically the two following ideas coming from Gelfand and Grothendieck:

1. Commutative rings correspond to geometric objects via functors

$$\text{commutative rings} \underset{\mathcal{O}}{\overset{\text{Spec}}{\rightleftarrows}} \text{geometric objects}$$

   In this vague setting we should think of Spec as in taking the spectrum of some collection of simultaneously diagonalized operators coming from the ring and $\mathcal{O}$ as returning the space of functions on these geometric objects. Again we should think of R as being the space of functions on $\operatorname{Spec} R$ in this correspondence.

2. Modules over rings R correspond to sheaves on the corresponding geometric object $\operatorname{Spec} R$ via

$$\text{modules} \underset{\text{global sections}}{\overset{\text{spectral decomposition}}{\rightleftarrows}} \text{sheaves}$$

   In particular in the previous examples we have that taking global sections amounts to taking the direct sum of modules, but this can also appear as a direct integral of modules in continuous versions of the previous examples. Spectral decomposition as we have seen is to break up a module into submodules where the ring action is pointwise multiplication.

We look at some examples of part 1. of the above philosophy.

Grothendieck's version of this idea is the heart of modern algebraic geometry. One correspondence is

$$\text{commutative rings} \longleftrightarrow \text{affine schemes}$$

but an earlier version might have been

$$\text{finitely presented, reduced, etc. } \mathbb{C}\text{-algebras} \longleftrightarrow \text{complex affine varieties}$$

In both correspondences, the Spec functor is given by taking the set of prime ideals. In the top correspondence, $\mathcal{O}$ is taking the structure sheaf of a scheme, but this amounts to taking polynomial functions on a space in the bottom correspondence.

Gelfand's version of this idea is the correspondence

$$\text{commutative } C^*\text{-algebras} \longleftrightarrow \text{Hausdorff locally compact topological spaces}$$

The functor Spec in this case is the eponymous Gelfand spectrum and $\mathcal{O}$ takes the continuous compactly supported functions on Hausdorff, locally compact spaces.

A special case of the above is the correspondence

$$\text{commutative von Neumann algebras} \longleftrightarrow \text{measure spaces}$$

One direction is some kind of spectrum, but the other direction is taking L^∞ functions on a measure space. As a side remark, there are only five commutative von Neumann algebras up to equivalence; they are L^∞ of: finite sets, $\mathbb{N}$, $[0,1]$, $[0,1]$ union a finite set, and $[0,1]$ union a countably infinite set.

There is a version of part 2. for each of the above examples involving modules and sheaves, but we will not discuss them here, aside from mentioning that we can talk about algebraic, continuous, or measurable families of vector spaces (sheaves) in the various settings above.

The group algebra (over $\mathbb{C}$)

Let G be any group and (V,ρ) any complex representation of G (the below discussion would work for other fields instead of $\mathbb{C}$). Recall that $\mathbb{C}[G]$ is the unique object for which the diagram

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, hook] & \mathrm{Aut}(V) \arrow[r, hook] & \mathrm{End}(V) \\ \mathbb{C}[G] \arrow[urr, dashed, "\overline{\rho}"'] & & \end{tikzcd}$$

commutes; moreover, we think of $\mathbb{C}[G]$ as a "free" $\mathbb{C}$-algebra on G. This is because the functor taking G to $\mathbb{C}[G]$ is the left adjoint to the forgetful functor from $\mathbb{C}$-**alg** to **Group** (given by taking the group of units).

An explicit description of $\mathbb{C}[G]$ is the $\mathbb{C}$-vector space generated by the elements of G with multiplication induced by the products in $\mathbb{C}, G$. An alternative description of $\mathbb{C}[G]$ is the set of finitely supported functions from G to $\mathbb{C}$, with multiplication given by convolution:

$$(f*h)(g) = \sum_{\substack{(x,y)\\ xy=g}} f(x)h(y) = \sum_x f(x)h(x^{-1}g) \quad \text{(defined since } f,h \text{ are finitely supported)}$$

This product is no different from the product defined in the first description of $\mathbb{C}[G]$.

What is really happening is that we can take f,h from above and form the box product $f \boxtimes h \colon G \times G \to \mathbb{C}$, which is just the map $(x,y) \mapsto f(x)h(y)$. We want to form a map from G to $\mathbb{C}$, so consider $G \times G \xrightarrow{m} G$, the product in G. The pushforward of $f \boxtimes h$ via m is $f*h$:

$$m_*(f \boxtimes h) \quad \text{also denoted} \quad \int_m f \boxtimes h = \sum_{\substack{(x,y)\\ xy=-}} f(x)h(y) = f*h$$

An important observation to make is that complex representations of G coincide with $\mathbb{C}[G]$-modules.

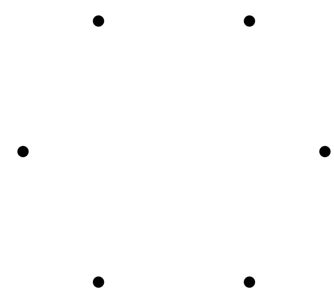
Representation theory of finite Abelian groups and the dual group

If G is Abelian, then $\mathbb{C}[G]$ is a commutative ring. So representation theory may be thought of as a special case of the study of modules over commutative rings. Using language from earlier, note that the corresponding geometric object to $\mathbb{C}[G]$ is $\operatorname{Spec}(\mathbb{C}[G])$, and representations of G correspond to vector bundles or sheaves over $\operatorname{Spec}(\mathbb{C}[G])$.

Let G be a finite Abelian group, and let $\widehat{G} \coloneqq \operatorname{Spec}(\mathbb{C}[G])$; we call this the dual of G. Then $\mathbb{C}[G]$ (functions on G) with multiplication given by convolution is isomorphic as an algebra to $\mathbb{C}[\widehat{G}]$ (functions on $\widehat{G}$, which we will see is a finite set) with pointwise multiplication. This isomorphism is usually given by some kind of finite Fourier transform. On the other hand, representations of G correspond to vector bundles (sheaves) on $\widehat{G}$, and this is some kind of finite spectral theorem.

Let $G = \langle x \rangle$ be a cyclic group of order n. The dual object $\widehat{G}$ is given by the n-th roots of unity in $\mathbb{C}$: (the picture is for $n = 6$)

$$\widehat{G} = \operatorname{Spec}(\mathbb{C}[G]) =$$



In this case, Spec is taking the maximal ideals of rings. With $\mathbb{C}[G] \cong \mathbb{C}[x]/(x^n - 1)$, it is equivalent to describe the elements in $\operatorname{Hom}_{\mathbb{C}\textbf{-alg}}(\mathbb{C}[x]/(x^n-1), \mathbb{C})$, which are characterized by each of the n-th roots of unity.

We take a small detour to discuss Schur's lemma, which states that for any group G, $\mathbb{C}[G]$-module homomorphisms between simple $\mathbb{C}[G]$-modules are either 0 or are isomorphisms. This follows by investigating kernels and images since they are submodules of V, W respectively. In particular, any $\mathbb{C}[G]$-module endomorphism of a simple $\mathbb{C}[G]$-module is a scalar multiple of the identity, since for any nonzero endomorphism T we can consider $T - \lambda \operatorname{id}_V$, which is no longer an isomorphism and hence must be zero. Here we used the algebraic closedness of $\mathbb{C}$, and in fact we could have replaced $\mathbb{C}$ by any algebraically closed field k.

Schur's lemma is used to prove that if G is Abelian then any finite-dimensional irrep of G is one-dimensional. If such an irrep V had finite dimension greater than or equal to 2, then left multiplication $V \xrightarrow{g} V$ by any element $g \in G$ must be a scalar multiple of id_V, which contradicts the irreducibility of V (since we assumed $\dim V \geq 2$).

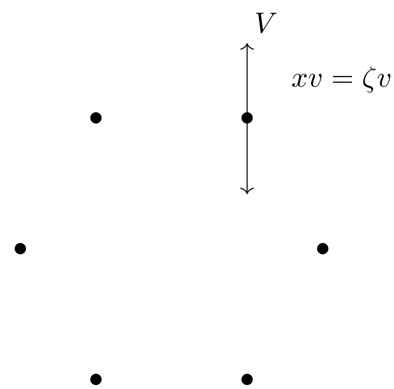
The action of G on the irrep V is given by scalar multiplication by $\chi_V(g) \in \mathbb{C}^\times$ since V is one-dimensional. More importantly, the assignment $g \mapsto \chi_V(g)$ is a group homomorphism $G \to \mathbb{C}^\times$, called a character of G.

For G Abelian, $\widehat{G}$ is the set of maximal ideals of $\mathbb{C}[G]$, which is in bijection with $\operatorname{Hom}_{\mathbb{C}\textbf{-alg}}(\mathbb{C}[G], \mathbb{C})$. By the universal property of the group algebra, $\mathrm{Hom}_{\mathbb{C}\text{-}\mathbf{alg}}(\mathbb{C}[G], \mathbb{C})$ is in bijection with $\mathrm{Hom}_{\mathbf{Group}}(G, \mathbb{C}^\times)$:

$$\begin{tikzcd}
G \arrow[r, "f"] \arrow[d, hook] & \mathrm{Aut}(\mathbb{C}) \cong \mathbb{C}^\times \arrow[r, hook] & \mathrm{End}(\mathbb{C}) \cong \mathbb{C} \\
\mathbb{C}[G] \arrow[urr, dashed, "\overline{f}"'] & &
\end{tikzcd}$$

Every map $\overline{f} \in \mathrm{Hom}_{\mathbb{C}\text{-}\mathbf{alg}}(\mathbb{C}[G], \mathbb{C})$ corresponds uniquely to a map $f \in \mathrm{Hom}_{\mathbf{Group}}(G, \mathbb{C}^\times)$. So in general we can also think of $\widehat{G}$ as the collection of irreps of G.

Given an irrep V of G (i.e. a simple $\mathbb{C}[G]$-module), we can try to form a sheaf out of V on $\widehat{G}$. Since V is irreducible, this sheaf has to be concentrated at only one of the points of $\widehat{G}$, and this point is the point corresponding to the irrep V itself. So in the example where $G = \langle x \rangle$ has order n, the point at which an irrep V lies on is the root of unity ζ for which the action of x on an irrep V is given by multiplication by ζ.



Since $\widehat{G}$ is in bijection with $\mathrm{Hom}_{\mathbf{Group}}(G, \mathbb{C}^\times)$, we can equip $\widehat{G}$ with a group operation; by doing so, G and $\widehat{G}$ are non-canonically isomorphic as groups. So in particular finite cyclic groups are (non-canonically) self-dual.

Next time, we will explore the Fourier transform, which for G finite Abelian is a $\mathbb{C}[G]$-module isomorphism

$$\mathrm{Fun}(G) \xrightarrow{\widehat{\ }} \mathrm{Fun}(\widehat{G})$$

where $\mathrm{Fun}(G)$, $\mathrm{Fun}(\widehat{G})$ are function spaces on $G, \widehat{G}$ and are given convolution and pointwise multiplication, respectively. Note that in this case $\mathrm{Fun}(G) \cong \mathbb{C}[G]$, so we give $\mathrm{Fun}(G)$ the corresponding action, which is given by $gf(x) = f(g^{-1}x)$. We will look at the action of G on $\mathrm{Fun}(\widehat{G}) \cong \mathbb{C}[\widehat{G}]$ next time. This isomorphism has some symmetry which is part of the statement of Pontryagin duality.

Lecture 3 September 02

More on the structure of $\widehat{G}$

For G a finite Abelian group, we saw/will see four equivalent descriptions of $\widehat{G}$:

$$\begin{aligned}\widehat{G} &= \text{set of irreps of } G\\ &= \mathrm{Hom}_{\mathbf{Group}}(G, \mathbb{C}^\times)\\ &= \text{unitary irreps of } G = \mathrm{Hom}_{\mathbf{Group}}(G, \mathrm{U}(1))\\ &= \mathrm{Spec}(\mathbb{C}[G]) = \text{maximal ideals of } \mathbb{C}[G] = \mathrm{Hom}_{\mathbb{C}\textbf{-alg}}(\mathbb{C}[G], \mathbb{C})\end{aligned}$$

Some of these descriptions for $\widehat{G}$ stop becoming equivalent to each other if we remove the adjectives finite or Abelian from G.

For any commutative ring R, recall that we think of R as functions on $\mathrm{Spec}(R)$ with pointwise multiplication. In general there is no reason to expect $\mathrm{Spec}(R)$ to be a group, but we were able to give $\mathrm{Spec}(\mathbb{C}[G])$ a group operation when G is Abelian. So in this situation something special is happening.

The deeper lesson is that the group algebra $\mathbb{C}[G]$ is distinguished among $\mathbb{C}$-algebras: Suppose now that G is any group (not necessarily Abelian). If V, W are representations of G then we can form a new representation $V \otimes_{\mathbb{C}} W$ where $g(v \otimes w) = gv \otimes gw$. This comes from the diagonal embedding $\mathbb{C}[G] \xrightarrow{\Delta} \mathbb{C}[G] \otimes_{\mathbb{C}} \mathbb{C}[G]$ given by $\mathbb{C}$-linearly extending the assignment $g \mapsto g \otimes g$. By doing so Δ is a $\mathbb{C}$-algebra homomorphism; it suffices only to check that it respects the group multiplication:

$$\Delta(hg) = hg \otimes hg = (h \otimes h)(g \otimes g) = \Delta(h)\Delta(g)$$

The natural action of $\mathbb{C}[G] \otimes_{\mathbb{C}} \mathbb{C}[G]$ on $V \otimes_{\mathbb{C}} W$ given by $(f \otimes g)(v \otimes w) = fv \otimes gw$ can be pulled back along the diagonal embedding Δ to give $V \otimes_{\mathbb{C}} W$ the action of G mentioned above.

If A is a $\mathbb{C}$-algebra and V, W are A-modules, there is a natural action of $A \otimes_{\mathbb{C}} A$ on $V \otimes_{\mathbb{C}} W$ given by $(a \otimes a')(v \otimes w) = av \otimes a'w$. But the diagonal map $A \to A \otimes_{\mathbb{C}} A$ might not be a $\mathbb{C}$-algebra homomorphism, but only $\mathbb{C}$-linear. In this case there may not be a natural way for A to act on $V \otimes_{\mathbb{C}} W$ like in the case of the group algebra. That the diagonal map $\mathbb{C}[G] \xrightarrow{\Delta} \mathbb{C}[G] \otimes_{\mathbb{C}} \mathbb{C}[G]$ for the group algebra is a $\mathbb{C}$-algebra homomorphism is part of the group algebra really being a Hopf algebra (see Wikipedia), and in this setting Δ is the comultiplication map.

If V is a representation of G, another representation of G we can form is the dual space $V^* := \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$, where the action of G on $\mathbb{C}$ is the trivial action. The action of G on V^* is $(gf)(v) = f(g^{-1}v)$. We may not be able to construct dual modules for an arbitrary $\mathbb{C}$-algebra A and A-module V, since we need some way to invert the action of A when "passing the action inside the function". This is related to one other datum that Hopf algebras have, which is their coinverse (also called antipode) map. In the case of the group algebra $\mathbb{C}[G]$, the coinverse map is $g \mapsto g^{-1}$ and this actually gives an isomorphism of $\mathbb{C}[G]$ and its opposite ring $\mathbb{C}[G]^{\mathrm{op}}$ (the ring where $g \cdot_{\mathrm{op}} h = hg$).

If A is a Hopf algebra over $\mathbb{C}$ and V, W are A-modules it is possible to define actions of A on $V \otimes_{\mathbb{C}} W$ and on $V^* := \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$ in a manner analogous to the above by using the comultiplication and antipode maps on A, respectively.

Denote by $\mathbf{Rep}(G)$ the category of finite-dimensional complex representations of G. The tensor product of representations gives $\mathbf{Rep}(G)$ a monoidal structure. Since we can also define duals of representations, $\mathbf{Rep}(G)$ is a rigid symmetric monoidal category; see the nLab.

Returning to G finite Abelian, in the description of $\widehat{G}$ as the collection of irreps of G, the group operation in $\widehat{G}$ is given by the tensor product of representations and inverses are given by taking dual spaces. Of course, the tensor product is associative and commutative in this case, and observe that $V \otimes_{\mathbb{C}} V^* \cong \mathrm{End}(V) \cong \mathbb{C}$ (since V is one-dimensional). The action of G on $V \otimes_{\mathbb{C}} V^*$ is also trivial: $g(v \otimes f) = gv \otimes gf = \chi_V(g) v \otimes \chi_V^{-1}(g) f = v \otimes f$.

The Fourier transform on finite Abelian groups

Let G be a finite Abelian group. Then $G \times \widehat{G}$ has a distinguished complex-valued function $G \times \widehat{G} \xrightarrow{\chi} \mathbb{C}^\times$ called the universal character, defined by

$$\chi(g, \hat{g}) = \chi_{\hat{g}}(g) \quad (= \hat{g}(g) \text{ if we think of } \hat{g} \text{ as a character})$$

where $\chi_{\hat{g}}$ is the character corresponding to $\hat{g} \in \widehat{G}$ (here $\widehat{G}$ is thought of as a set of points). Observe that χ is multiplicative in each component; that is,

$$\chi(hg, \hat{h}\hat{g}) = \chi_{\hat{h}\hat{g}}(hg) = \chi_{\hat{h}\hat{g}}(h)\chi_{\hat{h}\hat{g}}(g) = \chi_{\hat{h}}(h)\chi_{\hat{g}}(h)\chi_{\hat{h}}(g)\chi_{\hat{g}}(g)$$

For each $g \in G$, the function $\chi_g := \chi(g, -) \colon \widehat{G} \to \mathbb{C}^\times$ is a group homomorphism:

$$\chi_g(\hat{h}\hat{g}) = \chi(g, \hat{h}\hat{g}) = \chi_{\hat{h}}(g)\chi_{\hat{g}}(g) = \chi_g(\hat{h})\chi_g(\hat{g})$$

The assignment $g \mapsto \chi_g$ is a group homomorphism $G \xrightarrow{\chi_-} \mathrm{Hom}_{\mathbf{Group}}(\widehat{G}, \mathbb{C}^\times)$ that informally speaking, takes an element g to the map $\widehat{G} \to \mathbb{C}^\times$ which evaluates a character at g:

$$\chi_{gh}(\hat{g}) = \chi(gh, \hat{g}) = \chi(g, \hat{g})\chi(h, \hat{g}) = \chi_g(\hat{g})\chi_h(\hat{g}) \quad \text{Thus } \chi_{gh} = \chi_g \cdot \chi_h \text{ (pointwise multiplication)}$$

Denote by $\widehat{\widehat{G}}$ the group $\mathrm{Hom}_{\mathbf{Group}}(\widehat{G}, \mathbb{C}^\times)$. Since the kernel of χ_- is trivial and the image of χ_- is all of $\widehat{\widehat{G}}$, we obtain a (canonical) isomorphism of G with $\widehat{\widehat{G}}$. This is the content of Pontryagin duality for finite Abelian groups. The same result is true for locally compact Abelian groups, but requires some more tools to prove.

An analyst cannot help but think of the universal character $\chi(-, -)$ as being similar to kernels of integral operators (which are usually denoted $K(-, -)$ for example), or in this finite setting it might be more accurate to think of the universal character as similar to a matrix. The similarity is no coincidence. The universal character $\chi(-, -)$ is the kernel for the Fourier transform.

The Fourier transform is a map $\mathrm{Fun}(G) \xrightarrow{\widehat{-}} \mathrm{Fun}(\widehat{G})$ which we think of as a linear transformation with matrix $\chi(-,-)$. The Fourier transform is given by the formula

$$\hat{f}(\hat{g}) = \sum_{g\in G} \chi(g,\hat{g}) f(g)$$

For infinite groups that admit a Fourier transform like this, the sum is replaced by an integral of some kind and the kind of function spaces we consider may need adjustment.

One other perspective of this map is that it comes from a pullback and a pushforward. Consider the diagram

$$\begin{tikzcd} & G \times \widehat{G} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ G & & \widehat{G} \end{tikzcd}$$

and consider a function $f \in \mathrm{Fun}(G)$. The Fourier transform $\hat{f}$ is given by pulling back f along π_1, multiplying by $\chi(-,-)$, and then pushing forward along π_2. That is,

$$\hat{f} = \pi_{2*}(\pi_1^* f \cdot \chi)$$

The pushforward just means to sum (integrate) over fibers, and the fiber of $\hat{g}$ is $\{(g,\hat{g}) \mid g \in G\}$. Pulling back f along this projection does nothing but view f as a function on $G \times \widehat{G}$; that is, $(\pi_1^* f)(g,\hat{g}) = f(g)$. This recovers the first formula above for the Fourier transform.

As a fun fact, we can think about matrix multiplication this way. An $n \times m$ matrix $A = (A_{ij})$ is a function on the set $\{1,\dots,m\} \times \{1,\dots,n\}$ and a vector v in $\mathbb{C}^m$ is a function on $\{1,\dots,m\}$. There is a diagram

$$\begin{tikzcd} & \{1,\dots,m\} \times \{1,\dots,n\} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \{1,\dots,m\} & & \{1,\dots,n\} \end{tikzcd}$$

So Av coincides with $\pi_{2*}(\pi_1^* v \cdot A)$; that is, $(Av)_j = \sum_{i=1}^m A_{ij} v_i$ as expected. The same thing can be done for integral operators with kernel $K(-,-)$.

Denote by δ_g the Kronecker delta or the delta function(al) at g, or also the indicator function of $\{g\}$:

$$\delta_g(h) = \delta_{gh} = \begin{cases} 1 & \text{if } g = h \\ 0 & \text{otherwise} \end{cases}$$

Since G is finite, these delta functions span $\mathrm{Fun}(G)$. The Fourier transform of the delta function is calculated by

$$\widehat{\delta_g}(\hat{g}) = \sum_{h\in G} \chi(h,\hat{g}) \delta_g(h) = \chi(g,\hat{g}), \quad \text{so } \widehat{\delta_g} = \chi_g$$

which we summarize by saying that Fourier transforms of delta functions are characters.

From the calculation

$$(\delta_g * \delta_h)(x) = \sum_y \delta_g(y)\delta_h(y^{-1}x) = \delta_h(g^{-1}x) = \begin{cases} 1 & \text{if } x = gh \\ 0 & \text{otherwise} \end{cases}$$

we have $\delta_g * \delta_h = \delta_{gh}$. Recalling that $\chi_{gh} = \chi_g \cdot \chi_h$, conclude from the linearity of convolution that for any $f, h \in \mathrm{Fun}(G)$ that $\widehat{f * h} = \hat{f} \cdot \hat{h}$, which is summarized by saying that the Fourier transform turns convolution to pointwise multiplication. In other words, the Fourier transform diagonalizes the action of G on $\mathrm{Fun}(G)$ since $gf = \delta_g * f$. So in order for the Fourier transform to be a $\mathbb{C}[G]$-module homomorphism, G should act on $\mathrm{Fun}(\widehat{G})$ by pointwise multiplication by χ_g; that is, $gF(x) = \chi_g(x)F(x)$ for $F \in \mathrm{Fun}(\widehat{G})$.

Fourier inversion on finite Abelian groups

We will show that the Fourier transform has a G-linear inverse and hence is an isomorphism.

From some of the above calculations, we have that $g\delta_h = \delta_g * \delta_h = \delta_{gh}$. This equation shows that the action of G on the delta functions is by pushforward via left multiplication. So in reality the delta function in this case is a distribution or measure as opposed to a function.

Let G be any finite group. The left multiplication of G on itself gives rise to a natural right action on functions by pullback along the left multiplication map and a natural left action on distributions (or measures) by pushforward via the right action map on $\mathrm{Fun}(G)$. That is, given a complex-valued function $G \xrightarrow{f} \mathbb{C}$, the right action of G on f is given by $fg = (g\cdot)^*f = f(g-)$:

$$\begin{tikzcd} G \arrow[r, "g\cdot"] \arrow[dr, "f(g-)=(g\cdot)^*f"'] & G \arrow[d, "f"] \\ & \mathbb{C} \end{tikzcd} \quad \text{pulling back } f \text{ along } g\cdot$$

The right action on $\mathrm{Fun}(G)$ defines a map $\mathrm{Fun}(G) \xrightarrow{\cdot g} \mathrm{Fun}(G)$, by which we will pushforward distributions. Given a distribution on G; that is, a linear function $\mathrm{Fun}(G) \xrightarrow{T} \mathbb{C}$, the left action of G on T is given by $gT = (\cdot g)_*T = T(-g^{-1})$:

$$\begin{tikzcd} \mathrm{Fun}(G) \arrow[r, "\cdot g"] \arrow[d, "T"'] & \mathrm{Fun}(G) \arrow[dl, "(\cdot g)_*T=T(-g^{-1})"] \\ \mathbb{C} & \end{tikzcd} \quad \text{pushing forward } T \text{ by } \cdot g$$

Here, $(\cdot g)_*T = T(-g^{-1})$ because

$$(\cdot g)_*T(f) = \sum_{\substack{h \\ hg=f}} T(h) = T(fg^{-1}) \quad \text{(only one element in the fiber)}$$

A distribution in this setting is a continuous linear functional on $\mathrm{Fun}(G)$; where $\mathrm{Fun}(G) \cong \mathbb{C}^{|G|}$ is given the norm topology For other groups, $\mathrm{Fun}(G)$ is replaced by some other function space with a different topology. Typically, for a function $f\colon X \to Y$ and a distribution $F\colon \operatorname{Fun}(X) \to \mathbb{C}$, the pushforward $(f^*-)_*F$ is denoted by f_*F, and this is what is meant by the pushforward of the distribution F by f (similarly for measures).

It is instructive to look at the left action of G on the regular distributions, which are the distributions T_f obtained by integration against $f \in \operatorname{Fun}(G)$:

$$T_f(h) = \sum_{g\in G} f(g)h(g)$$

and the action of G on a regular distribution looks like

$$\begin{aligned}(gT_f)(h) = ((\cdot g)_*T_f)(h) = T_f(hg) = T_f(h(g-))\\ = \sum_{x\in G} f(x)h(gx) = \sum_{x\in G} f(g^{-1}x)h(x) = \sum_{x\in G} (fg^{-1})(x)h(x) = T_{fg^{-1}}(h)\end{aligned}$$

To summarize, for any $f \in \operatorname{Fun}(G)$, $gT_f = T_{fg^{-1}}$. For finite groups, every distribution on G is a regular distribution. Given a distribution T on G, define $f \in \operatorname{Fun}(G)$ by $f(g) = T(\delta_g)$. Then by linearity of T the calculation

$$T_f(\delta_h) = \sum_{x\in G} f(x)\delta_h(x) = \sum_{x\in G} T(\delta_x)\delta_h(x) = T(\delta_h)$$

implies that $T = T_f$ (and that f is unique).

Compare the original left action of G (not the natural one defined by pullbacks above) on the function δ_h with the left action of G on the regular distribution T_{δ_h}:

$$g\delta_h = \delta_{gh} \quad \text{and} \quad gT_{\delta_h} = T_{\delta_h g^{-1}} = T_{\delta_{gh}}$$

So this is a reason for why we mentioned above that the delta function was really a distribution or measure. The original definition of the left action of G on $\operatorname{Fun}(G)$ may be thought of as turning the natural right action of G on $\operatorname{Fun}(G)$ into a left action by being clever with signs, or more interestingly as identifying all functions with their corresponding regular distributions.

Now return to the case when G is finite Abelian. View $\chi_{\hat{g}}$ as an element of $\operatorname{Fun}(G)$. The calculation

$$g\chi_{\hat{g}}(h) = \chi_{\hat{g}}(g^{-1}h) = \chi_{\hat{g}}(g^{-1})\chi_{\hat{g}}(h) = \chi_{\hat{g}}(g)^{-1}\chi_{\hat{g}}(h) = \overline{\chi_{\hat{g}}(g)}\chi_{\hat{g}}(h)$$

shows that under the action of G on $\operatorname{Fun}(G)$, $\chi_{\hat{g}}$ is an eigenvector with eigenvalue $\chi_{\hat{g}}(-)^{-1} = \overline{\chi_{\hat{g}}(-)}$. In the theory of Fourier series (which we will see more later), a character of $\mathrm{U}(1) = \mathbb{R}/\mathbb{Z}$ is of the form $x \mapsto \exp(2\pi inx)$. The action of $\mathbb{R}/\mathbb{Z}$ on $\operatorname{Fun}(\mathbb{R}/\mathbb{Z})$ is given by $(x \cdot f)(\theta) = f(\theta - x)$, so

$$y \cdot \exp(2\pi inx) = \exp(2\pi in(x-y)) = \overline{\exp(2\pi iny)}\exp(2\pi inx)$$

(See Terence Tao's blog.)

Earlier we showed that $\widehat{\delta_g} = \chi_g$, and that extending this linearly produces the Fourier transform. We would like for the inverse Fourier transform to send $\delta_{\hat{g}}$, the indicator function of $\hat{g}$, to a character, to match the Fourier transform.

(Note that the space $\mathrm{Fun}(\widehat{G})$ has basis the indicator functions $\delta_{\hat{g}}$.) But the inverse Fourier transform should also respect the group action on the function spaces. Compared to the action of G on $\mathrm{Fun}(G)$, the action of G on $\mathrm{Fun}(\widehat{G})$ is given by identifying G with $\widehat{\widehat{G}} \subset \mathrm{Fun}(\widehat{G})$ and using pointwise multiplication. That is, $gf = \chi_g \cdot f$. We should try another calculation to see if that provides a clue; in particular, if we want to have a chance at defining the inverse Fourier transform we should try to take Fourier transforms of characters. The Fourier transform of $\chi_{\hat{g}} \in \mathrm{Fun}(G)$ is given by

$$\widehat{\chi_{\hat{g}}}(\hat{h}) = \sum_{h \in G} \chi(h, \hat{h}) \chi_{\hat{g}}(h) = \sum_{h \in G} \chi(h, \hat{h}\hat{g}) = |G| \delta_{\hat{g}^{-1}},$$

where in the last equality we used the following result:

$$\sum_{g \in G} \chi_g(\hat{h}) = \begin{cases} 0 & \text{if } \hat{h} \neq 1_{\widehat{G}} \\ |G| & \text{otherwise} \end{cases}$$

Indeed, since $\widehat{\widehat{G}}$ is a group, if $\hat{h} \neq 1_{\widehat{G}}$, then choose $h \in G$ for which $\chi_{\hat{h}}(h) = \chi_h(\hat{h}) \neq 1$ (we can do this since $\chi_{\hat{h}}$ is not the trivial character). Then with $\chi_h \in \widehat{\widehat{G}}$, we have

$$\chi_h(\hat{h}) \sum_{g \in G} \chi_g(\hat{h}) = \sum_{g \in G} \chi_{hg}(\hat{h}) = \sum_{g \in G} \chi_g(\hat{h}),$$

but $\chi_h(\hat{h}) \neq 1$, so $\sum_{g \in G} \chi_g(\hat{h}) = 0$. If $\hat{h} = 1_{\widehat{G}}$, then $\sum_{g \in G} \chi_g(\hat{h}) = \sum_{g \in G} 1 = |G|$. A similar argument may be done to evaluate the sum $\sum_{\hat{g} \in \widehat{G}} \chi_{\hat{g}}(h)$.

Based on the calculations above, one might try to calculate the Fourier transform of $\overline{\chi_{\hat{g}}}$, to obtain $|G|\delta_{\hat{g}}$. So the inverse Fourier transform should send $\delta_{\hat{g}}$ to $\chi_{\hat{g}}/|G|$, and by extending linearly we find that the inverse Fourier transform $\mathrm{Fun}(\widehat{G}) \xrightarrow{-^\vee} \mathrm{Fun}(G)$ is given by

$$F^\vee(g) = \frac{1}{|G|} \sum_{\hat{g} \in \widehat{G}} \overline{\chi(g, \hat{g})} F(\hat{g})$$

One can calculate that for any $f \in \mathrm{Fun}(G)$ that

$$f(g) = \frac{1}{|G|} \sum_{\hat{g} \in \widehat{G}} \overline{\chi(g, \hat{g})} \hat{f}(\hat{g})$$

which is the Fourier inversion theorem for finite Abelian groups.

By endowing $\mathrm{Fun}(G)$ with the Hermitian L^2 inner product

$$\langle f, h \rangle = \sum_{g \in G} f(g) \overline{h(g)}$$

it turns out that the characters $\chi_{\hat{g}}$ form an orthonormal basis (use one of the earlier calculations to do this). A similar statement is true for $\mathrm{Fun}(\widehat{G})$, and by scaling the Fourier transform, the inverse Fourier transform, and possibly the inner products correctly, we obtain an unitary isomorphism of the Hilbert spaces $\mathrm{Fun}(G)$ and $\mathrm{Fun}(\widehat{G})$ (i.e., the inner product is preserved).

A preview for locally compact Abelian groups

Some examples of locally compact Abelian (LCA) groups are the finite Abelian groups, $\mathbb{Z}$, $\mathrm{U}(1)$, $\mathbb{R}$, $\mathbb{Z}_p$ (the p-adic integers), and $\mathbb{Q}_p$ (the p-adic numbers). In this setting $\widehat{G}$ is defined to be the set of unitary irreps, otherwise given by $\mathrm{Hom}_{\mathbf{Group}}(G, \mathrm{U}(1))$.

Pontryagin duality still holds in this setting. The result encapsulates the following statements:

1. The canonical map $G \to \widehat{\widehat{G}}$ (where $\widehat{\widehat{G}} = \mathrm{Hom}_{\mathbf{Group}}(\widehat{G}, \mathrm{U}(1))$) is an isomorphism.
2. The Fourier transform and its inverse are Hilbert space isomorphisms (i.e. unitary isomorphisms) of $L^2(G)$ with $L^2(\widehat{G})$ (with the correct measures on G, $\widehat{G}$). The transforms interchange convolution with pointwise multiplication, and send delta "functions" to characters (neither of which are L^2 functions, so this is meant in the sense of distributions).
3. A spectral theorem: Representations of LCA groups are in correspondence with sheaves of vector spaces on $\widehat{G}$. Given a finite-dimensional representation V of G, decompose V into its invariant subspaces (otherwise called isotypic components) $V = \bigoplus_{\hat{g}} V_{\hat{g}}$ where $V_{\hat{g}} = \{v \in V \mid gv = \chi_{\hat{g}}(g)v\}$. The Fourier transform in this setting takes the representation (sheaf) $\mathrm{Fun}(G)$ and spits out the representation (sheaf) $\mathrm{Fun}(\widehat{G}) = \bigoplus_{\hat{g} \in \widehat{G}} \mathbb{C}$...

Soon we will look at the LCA groups $\mathrm{U}(1)$ and $\mathbb{R}$. There we can recover the usual Fourier series and real Fourier transform theory.

Lecture 4 September 04

Unitarizable representations and semisimplicity

A category is called semisimple if every object is the direct sum of finitely many simple objects. We will show that $\mathbf{Rep}(G)$ is semisimple for G finite.

A complex representation of a group G is unitarizable if there exists a Hermitian inner product (positive definite, sesquilinear bilinear form) $\langle -,-\rangle$ on V for which

$$\langle gv, gw\rangle = \langle v, w\rangle$$

In other words, V is unitarizable if it can be endowed with a Hermitian inner product that is invariant under the group action, or otherwise we can say that the group acts by unitary transformations on V with this inner product. In the language of diagrams, given the representation (V,ρ), there exists an inner product $\langle -,-\rangle$ and the map $G \xrightarrow{\rho} \mathrm{U}(V,\langle -,-\rangle)$ (with the same definition as $G \xrightarrow{\rho} \mathrm{GL}(V)$) for which the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[dr, dashed, "\rho"'] & \mathrm{GL}(V) \\ & \mathrm{U}(V,\langle -,-\rangle) \arrow[u, hook] \end{tikzcd}$$

A non-example of a unitarizable representation: If $\mathbb{Z}$ acts on a one-dimensional vector space V and $1 \in \mathbb{Z}$ acts by an operator outside of $\mathrm{U}(V)$, then there is no way to unitarize V. The key point here is that since V is one-dimensional, the definition of $\mathrm{U}(V)$ does not depend on any choice of Hermitian inner product on V.

Given a representation (V,ρ), recall we can define the dual representation V^* where $gf(v) = f(g^{-1}v)$. In this case we think of the action of G as being inverted since we use the inversion map on G to define the action on V^*. On the other hand, since the action of G on V is by complex-valued matrices, we can pointwise conjugate these matrices to obtain a representation $(\overline{V},\overline{\rho})$ where the underlying Abelian group $\overline{V}$ coincides with the Abelian group V but the scalar multiplication is given by $c \cdot_{\overline{V}} v = \overline{c} \cdot_V v$. Extending this, the group G acts on $\overline{V}$ by $gv = \overline{\rho(g)}v$, and in this case we think of the action as being conjugated as opposed to inverted in the case with the dual vector space.

A Hermitian inner product $\langle -,-\rangle$ on any complex vector space V corresponds to an isomorphism of $\overline{V}$ with V^* given by $v \mapsto \langle -, v\rangle$, and from any isomorphism $\overline{V} \xrightarrow{H} V^*$, define the inner product $\langle -,-\rangle$ by $\langle v, w\rangle = H(v)(w)$. The G-invariant inner products correspond to G-equivariant isomorphisms H.

That a representation V is unitarizable is a property of V as opposed to being part of the structure. On the other hand, if W is a unitary representation, part of the data of W is a fixed G-invariant inner product $\langle -,-\rangle$, as opposed to a unitarizable representation V for which we do not specify any one G-invariant inner product. Of course, once we do specify a G-invariant inner product $\langle -,-\rangle_0$, we can speak of the unitary representation $(V, \langle -,-\rangle_0)$, which is different from the unitary representation $(V, \langle -,-\rangle_1)$ for a different G-invariant inner product $\langle -,-\rangle_1$.

If V is a finite-dimensional unitarizable representation, then V is a semisimple representation; that is, V is the direct sum of irreps/simples. Assume V is already not simple, and fix a G-invariant inner product $\langle -,-\rangle$ on V. Let $W \subset V$ be any G-invariant subspace, and see that $W^\perp = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}$ is a G-invariant subspace of V since the W and the inner product $\langle -,-\rangle$ are G-invariant:

$$\langle gv, w\rangle = \langle v, g^{-1}w\rangle \quad \text{implies } gv \in W^\perp \text{ whenever } v \in W^\perp$$

Then induction proves the result (this amounts to iterating this procedure on $W^\perp$ and repeating on each of the resulting subspaces until it cannot be done, which happens because V is finite-dimensional).

Another nice result is that if G is finite, then a finite-dimensional representation V is unitarizable and hence semisimple. The heart of this result is due to the existence of an averaging element $\mathrm{av} \in \mathbb{C}[G]$ given by

$$\mathrm{av} = \frac{1}{|G|}\sum_{g\in G} g$$

Note that $\mathrm{av} \in Z(\mathbb{C}[G])$. To define this averaging element, it is crucial that G is finite and that the characteristic of the field $\mathbb{C}$ (zero) does not divide the order of the group. The existence of an averaging operator is the content of Maschke's theorem from earlier, since we can use this averaging element to form a projector needed to form complements of G-stable subspaces, as we will see.

Let V be any finite-dimensional representation. Then we can define the subspace of G-fixed points, V^G, by

$$V^G = \{v \in V \mid gv = v \text{ for all } g \in G\}$$

The subspace V^G is the part of V that G acts trivially on, or in other words is the trivial representation part of V when broken into irreps. For any $v \in V$, $\mathrm{av}\cdot v \in V^G$ since for any $h \in G$ we have

$$h(\mathrm{av}\cdot v) = h\left(\frac{1}{|G|}\sum_{g\in G} gv\right) = \left(\frac{1}{|G|}\sum_{g\in G} hgv\right) = \left(\frac{1}{|G|}\sum_{g\in G} gv\right) = \mathrm{av}\cdot v$$

Observe further that because the sum is normalized with the factor $1/|G|$, multiplication by av defines a projection map to V^G, since $\mathrm{av}^2 = \mathrm{av}$, so $\mathrm{av}-$ is an idempotent. The complementary idempotent is $(1-\mathrm{av})-$, and see that it squares to itself as well. It is complementary since $\mathrm{av}(1-\mathrm{av}) = 0$. It follows that $V = \mathrm{av}V \oplus (1-\mathrm{av})V = V^G \oplus (1-\mathrm{av})V$.

An aside: To prove Maschke's theorem, we do a similar trick. Given any G-invariant subspace H of a representation V, consider the space $\mathrm{Hom}_{\mathbb{C}}(V, H)$, which contains many projections since short exact sequences in $\mathbf{Vect}_{\mathbb{C}}$ always split, or this is just true from linear algebra. What we want to find is a projector π which is G-intertwining; that is, $(g\cdot)\pi(g^{-1}\cdot) = \pi$. Considering the action of G on $\mathrm{Hom}_{\mathbb{C}}(V, H)$ by conjugation, finding a G-intertwining projector

π amounts to finding a projector that is fixed under conjugation. Pick any one projector $V \xrightarrow{\pi_0} H$, and average π_0 using av to get the projector

$$\pi = \mathrm{av} \cdot \pi_0 = \frac{1}{|G|} \sum_{g \in G} (g \cdot) \pi_0 (g^{-1} \cdot)$$

Hence π is a fixed point under the conjugation action and thus is a desired G-equivariant projector.

Returning to showing that a finite-dimensional representation V is unitarizable and hence semisimple, it amounts to finding a G-invariant inner product $\langle -, - \rangle$. Pick any inner product $\langle -, - \rangle_0$ on V and make it G-invariant by averaging it via av. Define the inner product $\langle -, - \rangle$ by

$$\langle -, - \rangle = \text{“av}(\langle -, - \rangle_0)\text{”} = \frac{1}{|G|} \sum_{g \in G} \langle g-, g- \rangle_0$$

Observe that this new inner product is G-invariant as desired, and is really a Hermitian inner product because $\langle -, - \rangle_0$ was to begin with. (In other words, we have found an isomorphism $H \in \mathrm{Hom}_{\mathbb{C}}(\overline{V}, V^*)^G$ by applying the projection $\mathrm{Hom}_{\mathbb{C}}(\overline{V}, V^*) \to \mathrm{Hom}_{\mathbb{C}}(\overline{V}, V^*)^G$ obtained via averaging to any isomorphism $H_0 \in \mathrm{Hom}_{\mathbb{C}}(\overline{V}, V^*)$.) Since V is unitarizable, V is semisimple. This shows that $\mathbf{Rep}(G)$ is semisimple. The slickness of the argument above gives a compelling reason to use the group algebra $\mathbb{C}[G]$.

Distributions and measures

When G is finite, then the group algbera $\mathbb{C}[G]$ can be identified with the algebra $\mathrm{Fun}(G)$ with convolution, and we identify av with $\mathbf{1}_G/|G|$ ($\mathbf{1}_G$ is the indicator function on G) and identify $1_{\mathbb{C}[G]}$ with δ_{1_G}, and everything in between. If G is also Abelian, then $\mathrm{Fun}(G)$ with convolution is isomorphic to $\mathrm{Fun}(\widehat{G})$ with pointwise multiplication via the Fourier transform. A short calculation shows that the Fourier transform of δ_{1_G} is $\mathbf{1}_{\widehat{G}} = \sum_{\hat{g} \in \widehat{G}} \delta_{\hat{g}}$, which can be thought of as some version of Plancherel's theorem.

So if G acts on V, then we know that V decomposes as $V = \bigoplus_{\hat{g} \in \widehat{G}} V_{\hat{g}}$, where $V_{\hat{g}} = \delta_{\hat{g}} V$, where we think of the delta functions as being projectors onto subspaces of V for which G acts by multiplication by characters. The notation is suggestive: $g \cdot \delta_{\hat{g}} V = \chi_{\hat{g}}(g) \delta_{\hat{g}} V$.

When we consider groups G that are not finite, we may need to consider a smaller function space than all of $\mathrm{Fun}(G)$ in order for the theory to be nicer. In doing so, the functions δ_{1_G}, $\mathbf{1}_G/|G|$ cease to belong in these new function spaces or even in $\mathrm{Fun}(G)$ to begin with, and should be thought of as distributions or measures. For example, this will happen with $G = \mathrm{U}(1)$ or $\mathbb{R}$ since we might consider the function space $L^2(G)$ instead of $\mathrm{Fun}(G)$. For example, the function $\mathbf{1}_G/|G|$ corresponds to the regular distribution $1/\mu(G) \int_G - \mathrm{d}\mu$ for a correctly chosen measure μ on G, for G compact, and δ_{1_G} as the functional which evaluates functions at 1_G.

Because the convolution of two functions is given by an integral or otherwise a pushforward, secretly somehow what is actually happening is that we are pushing forward a particular distribution or measure. “You don't integrate functions, you integrate measures.” So this is more evidence that we should see if we can do an analysis of the representations of infinite Abelian groups using distributions instead of functions.

Let G be a locally compact Hausdorff topological group. Then Haar's theorem (see Wikipedia) states that there is a measure μ on the Borel subsets of G, called the Haar measure, which is:

1. Left-translation invariant under the left multiplication action of G; that is, the pushforward measure $(g\cdot)_*\mu$ agrees with μ.
2. Finite on compact sets of G.
3. Outer regular on Borel sets; that is, $\mu(S) = \inf\{\mu(U) \mid S \subseteq U \text{ with } U \text{ open}\}$ for S a Borel set.
4. Inner regular on open sets; that is, $\mu(U) = \sup\{\mu(K) \mid K \subseteq U \text{ with } K \text{ compact}\}$ for U an open set.
5. Unique up to positive scaling.

If G is compact then we can normalize the Haar measure on G to produce a unique Haar measure for which $\mu(G) = 1$. This measure is provided by av, viewed as a measure by the formula $\mathrm{av}(E) = \frac{1}{\mu(G)} \int_E 1 \mathrm{d}\mu$ where μ is any Haar measure, or as the regular distribution T_{av} where $T_{\mathrm{av}}(f) = \int_G f(g)\mathrm{av}(g)\mathrm{d}\mu = \frac{1}{\mu(G)} \int_G f(g)\mathrm{d}\mu$ where μ is any Haar measure. So for example, the normalized Haar measure on $\mathrm{U}(1)$ is $\mathrm{d}\theta/2\pi$.

On a manifold, to be able to integrate we need to be able to find a top differential form. So if G is a Lie group, consider the tangent space $\mathfrak{g} = T_G(1_G)$, which is the only one we need to consider since all other tangent spaces are obtained via translation of the tangent space at the identity. Then the space of top differential forms $\bigwedge^{\dim \mathfrak{g}} \mathfrak{g}^*$ is a one-dimensional vector space. The normalized Haar measure is the measure induced by the normalized top form in this vector space.

The requirement that G is locally compact probably comes from wanting to be able to approximate integrals using compact exhaustions. The Hausdorff requirement is a consequence of the following result of topological groups: For G a topological group, let H be the closure of the set $\{1_G\}$. Then H is a normal subgroup and the quotient G/H is the largest quotient of G that is Hausdorff. Furthermore, every continuous map of G to a Hausdorff space factors through the quotient G/H. So in our setting, we will be considering continuous maps from G to the complex numbers, so in this setting G is indistinguishable from its largest Hausdorff quotient. As a consequence, there is no loss in defining LCA groups to be locally compact, Abelian, and Hausdorff.

A fun fact is that if G is compact, then the Haar measure on G is automatically both left and right-translation invariant, even if G is not Abelian. For any Haar measure μ on G and $g \in G$, the pushforward of μ by right multiplication by g^{-1} produces another left-invariant Haar measure $(\cdot g)_*\mu$. So μ is some multiple of $(\cdot g)_*\mu$, denote this multiple by $\mathrm{mod}(g)$, called the modulus character of G. It turns out that this quantity does not depend on the choice of Haar measure μ we started with, and is also a group homomorphism $G \to \mathbb{R}_{>0}$, which measures how much the left and right-translation invariance of the Haar measure disagree. The modulus character is trivial if G is Abelian. For G compact, there are no compact subgroups of $\mathbb{R}_{>0}$ aside from $\{1\}$, so in this case the modulus character is also trivial, and hence Haar measures on compact groups are bi-invariant. In other words, the Haar measure is a $G \times G$-invariant measure on G (where $(g, h) \in G \times G$ acts by left translation by g and by right translation by h^{-1}). This is because the element $\mathrm{av} \in \mathbb{C}[G]$ is $G \times G$-invariant (where $(g, h) \in G \times G$ acts by left multiplication by g and by right multiplication by h^{-1}); this is stronger than just being invariant under conjugation by G.

A tiny preview of LCA groups

Let G be compact. Then finite-dimensional representations of G are semisimple. The proof follows the same ideas as in the finite case, but we replace all sums with integrals using the normalized Haar measure. Note that the normalization of the Haar measure is what will ensure that our desired projector actually is an idempotent as we did in the finite case.

For LCA groups, it is also true that finite-dimensional representations are semisimple, and from before we know the irreps are one-dimensional. If $G = \mathrm{U}(1) = S^1$, then we obtain the theory of Fourier series, since the dual group $\widehat{S^1}$ is isomorphic to $\mathbb{Z}$; each character of S^1 is of the form $z \mapsto z^n$ (or $x \mapsto \exp(2\pi i n x)$ if we think of S^1 as $\mathbb{R}/\mathbb{Z}$) for some integer n.

Lecture 5 September 09

Fourier series

Let $G = \mathrm{U}(1)$, which we know familiarly as the circle. Since G is compact, representations of G are semisimple, and since G is Abelian, irreps of G are one-dimensional. We could decide to stop with the theory here, since at this point we have classified all of the finite-dimensional representations using the dual group $\widehat{G}$ from before.

The dual group of the circle $\widehat{\mathrm{U}(1)}$ is $\mathbb{Z}$, where $n \in \mathbb{Z}$ corresponds to the character $z \mapsto z^n$, or if we think of $\mathrm{U}(1)$ as $\mathbb{R}/\mathbb{Z}$, the character $x \mapsto \exp(2\pi i n x)$. We will think of the circle as $\mathbb{R}/\mathbb{Z}$ usually, to appease the analysts (this is a joke). In this setting, the universal character $\chi(-,-)$ is the map $(x, n) \mapsto \exp(2\pi i n x)$ (so $\chi_x = \chi(x, -)$ and $\chi_n = \chi(-, n)$). Then the Fourier transform as we mentioned before is obtained from the maps in the diagram

$$\begin{tikzcd} & G \times \widehat{G} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ G & & \widehat{G} \end{tikzcd}$$

by pulling back $f \in \mathrm{Fun}(G)$ along π_1, multiplying by $\chi(-,-)$, and then pushing forward along π_2. That is,

$$\hat{f} = \pi_{2*}(\pi_1^* f \cdot \chi) \quad \text{so } \hat{f}(n) = \int_0^1 f(x) \exp(2\pi i n x) \mathrm{d}x$$

The Fourier inversion theorem in this setting is

$$f(x) \approx \sum_{n \in \mathbb{Z}} \hat{f}(n) \exp(-2\pi i n x)$$

where the $\approx$ means that we should worry about convergence. It ends up being the case that we get convergence in the L^2 norm, but more analysis is needed for better results.

Another point of view of the Fourier transform is to look at $G = \mathrm{U}(1) = \mathbb{R}/\mathbb{Z}$ acting on $L^2(G)$ by translation. The action is

$$y \cdot f =: \tau_y f = f((-) - y) \quad \text{for } y \in \mathbb{R}/\mathbb{Z}$$

The group action preserves each of the subspaces $\mathbb{C}\chi_n$ since like before, we have $\tau_y \chi_n(x) = \chi_{-n}(y)\chi_n(x)$. We can think of the Fourier transform in linear algebra terms, since $L^2(G)$ has the Hermitian inner product $\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)} \mathrm{d}x$. Some analysis shows that the characters χ_n form an orthonormal basis of $L^2(G)$, so the Fourier transform finds the components of any $f \in L^2(G)$ using the inner product:

$$\hat{f}(n) = \langle f, \chi_{-n} \rangle = \int_0^1 f(x) \exp(2\pi i n x) \mathrm{d}x$$

The inverse Fourier transform reassembles f from its projections onto each of the subspaces $\mathbb{C}\chi_n$, but as we noted before, we might need to worry about convergence:

$$f(x) \approx \sum_{n \in \mathbb{Z}} \langle f, \chi_{-n} \rangle \chi_{-n} = \sum_{n \in \mathbb{Z}} \hat{f}(n) \exp(-2\pi i n x)$$

Analysts are likely used to the characters being maps $x \mapsto \exp(inx)$ and the Fourier transform of $f \in L^2(\mathrm{U}(1))$ as $\frac{1}{2\pi}\int_0^{2\pi} f(x)\exp(-inx)\mathrm{d}x$; to obtain this form we can present $\mathrm{U}(1)$ not by $\mathbb{R}/\mathbb{Z}$ but as unimodular points in $\mathbb{C}$, with the Haar measure given by arclength measure, and by reindexing the characters by interchanging n with $-n$.

Function spaces

An algebraic point of view is to view the algebra spanned by all of the characters as the algebraic functions on the circle; that is,

$$\bigoplus_{n\in\mathbb{Z}} \mathbb{C}\chi_n \cong \mathbb{C}[z, z^{-1}] \quad \text{where } z = \exp(2\pi i x) \text{ since } x \in \mathbb{R}/\mathbb{Z}$$

The coordinate ring of the affine space $\mathbb{C}$ is $\mathbb{C}[z]$. By localizing the coordinate ring at the maximal ideal $m_0 = \{f \in \mathbb{C}[z] \mid f(0) = 0\} = (z)$, we obtain the algebraic functions on $\mathbb{C}^\times$, which agree with the algebraic functions on the circle if we restrict z to $z = \exp(2\pi i x)$ for $x \in \mathbb{R}/\mathbb{Z}$.

An analytic point of view is to see this space of algebraic functions as being dense in function spaces like $L^2(G)$, $L^1(G)$, $C^\infty(G)$, $C^\omega(G)$ (analytic functions), $C^{-\infty}(G)$ (distributions), or even $C^{-\omega}(G)$ (hyperfunctions?). Let us suppress the (G) for now. Each of these spaces has the algebraic functions above as an "algebraic core", which we understand the Fourier transform on. The question is how the action of the Fourier transform extends to the rest of the function spaces, and this requires analysis. The Fourier transform exchanges the following spaces (this is not to say the Fourier transform is always an isomorphism of these spaces!):

$$\begin{aligned} L^2 &\longleftrightarrow \ell^2 & L^1 &\longleftrightarrow c_0 \\ C^\infty &\longleftrightarrow \{f \to 0 \text{ faster than } 1/\text{polynomial}\} & C^\omega &\longleftrightarrow \{f \to 0 \text{ faster than } 1/\text{exponential}\} \\ C^{-\infty} &\longleftrightarrow \{f \to \infty \text{ polynomial order}\} & C^{-\omega} &\longleftrightarrow \{f \to \infty \text{ exponential order?}\} \end{aligned}$$

It is actually the case that the Fourier transform gives an isomorphism of L^2 with ℓ^2. Harish-Chandra studied similar function spaces for non-Abelian groups and found similar exchanges as above.

More than just an orthonormal basis, the algebraic core of characters is the basis in which the action of G is diagonalized via the Fourier transform. This is due to the property that the Fourier transform interchanges convolution of functions with pointwise multiplication of functions. In the setting of the circle group, this is to recast the action of translation as convolution against a delta distribution and to use the Fourier transform to obtain pointwise multiplication by a scalar:

$$\tau_y f = \delta_y * f \overset{\widehat{-}}{\longmapsto} \widehat{\delta_y * f} = \exp(2\pi i(-)y)\hat{f}.$$

If f is the character χ_n, then τ_y acts on f by multiplication by the scalar $\exp(-2\pi i n y)$, and this action commutes with the Fourier transform. In general we should worry about what functions it makes sense to convolve against; after all, we are convolving functions with distributions, so some analysis must be done. If we restrict to continuous functions, then the analysis of C^*-algebras will appear, for example.

Since U(1) is thought of as continuous as opposed to discrete, there is an "infinitesimal" action of U(1) on functions given by taking the derivative. The Fourier transform takes differentiation to multiplication by a multiple of the identity function:

$$\widehat{\frac{\mathrm{d}}{\mathrm{d}x}f}(n) = -2\pi i n \hat{f}$$

If f is the character χ_m, then the derivative of f is $2\pi i m f$. The Fourier transform of $2\pi i m f$ is $2\pi i m \delta_{-m}$, so the action of the derivative commutes with the Fourier transform.

What is meant by the "infinitesimal" action $\frac{\mathrm{d}}{\mathrm{d}x}$? From the action of $\mathrm{U}(1) = \mathbb{R}/\mathbb{Z}$, we define an action of its Lie algebra $\mathbb{R}$ (the tangent space at the identity 0 of $\mathbb{R}/\mathbb{Z}$) on a function by $y \cdot f = \frac{\mathrm{d}}{\mathrm{d}t}(\tau_{ty} f)|_{t=0}$. In this case, $\frac{\mathrm{d}}{\mathrm{d}x}$ agrees with the action of -1 on functions; that is, $\frac{\mathrm{d}}{\mathrm{d}x} f = \frac{\mathrm{d}}{\mathrm{d}t}(\tau_{-t} f)|_{t=0}$. If we want to insist on working in U(1) given by the unimodular complex numbers, then the Lie algebra of U(1) in this case is $i\mathbb{R}$. The action of $i\mathbb{R}$ on g (here g is a function of θ; if g is a function of z, put $z = \exp(i\theta)$) is given by $iy \cdot g = \frac{\mathrm{d}}{\mathrm{d}t}(\exp(iyt) \cdot g)|_{t=0}$. In this case $\frac{\mathrm{d}}{\mathrm{d}\theta}$ is given by the action of $-i$ on functions; that is, $\frac{\mathrm{d}}{\mathrm{d}\theta} g = \frac{\mathrm{d}}{\mathrm{d}t}(\exp(-it) \cdot g)|_{t=0}$.

Due to Pontryagin duality, we can mirror the above theory by considering taking Fourier transforms of functions on $\mathbb{Z}$ to get functions on $\widehat{\mathbb{Z}} = \mathrm{U}(1)$. The universal character $\chi(-,-)$ in this case is the same map (but we should flip the inputs) $(n, x) \mapsto \exp(2\pi i n x)$ Even though $\mathbb{Z}$ is a discrete group, we can still think about difference operators in place of derivatives (these are just linear combinations of differences of translation operators). As desired, the action of $\mathbb{Z}$ on a function is given by translation, and the Fourier transform exchanges this action with multiplication by a character. That is, for $n \in \mathbb{Z}$

$$n \cdot g =: \tau_n g = \delta_n * g = g((-) - n) \overset{\widehat{-}}{\longmapsto} \widehat{\delta_n * g} = \exp(2\pi i n(-))\hat{g}$$

As before, we should also expect difference operators to correspond to multiplier operators after taking the Fourier transform.

The Fourier transform on $\mathbb{Z}$, which we also denote by $\mathcal{F}_{\mathbb{Z}}$, is given by $\hat{g}(x) = \sum_{n\in\mathbb{Z}} g(n)\exp(2\pi i n x)$, which is different from the inverse Fourier transform of the Fourier transform $\mathcal{F}_{\mathrm{U}(1)}$ from U(1) to $\mathbb{Z}$. Each Fourier transform has order four. Note also that by composing the two Fourier series together, we get $\mathcal{F}_{\mathbb{Z}}\mathcal{F}_{\mathrm{U}(1)} f(x) = f(-x)$. Similarly, $\mathcal{F}_{\mathrm{U}(1)}\mathcal{F}_{\mathbb{Z}} g(n) = g(-n)$.

Even though we understand well how the Fourier transform acts on delta distributions and characters, there is the issue of figuring out how the Fourier transforms extend to larger function spaces. There should be a nice way to package the theory together nicely, in a way that is agnostic to the exact functions appearing in the larger spaces. A nice way to package everything together is via the spectral theorem for the group U(1).

The spectral theorem for U(1)

Let $\mathcal{H}$ be any unitary representation of U(1) (note that some of the function spaces we mentioned earlier are not Hilbert spaces!). Since we have diagonalized the action of U(1), it acts on $\mathcal{H}$ by compact operators (in this case the operators are approximated by finite-rank diagonal operators). By the spectral theorem for compact operators, we can find countably many eigenspaces of $\mathcal{H}$ whose direct sum is dense in $\mathcal{H}$. We find these eigenspaces explicitly using the characters χ_n.

The first part of this is to find analogues of the group algebra in this setting. On one hand, compactly supported functions on $G = \mathbb{R}/\mathbb{Z}$ would be a good starting place, but we can choose even larger spaces for which the product, convolution, still makes sense. For example, suitable candidates for a group algebra might be the space of continuous or even L^1 functions on G. For now, consider the continuous functions on G with convolution, denoted by $(C(G), *)$. A unitary representation $(\mathcal{H}, \rho)$ of G is in correspondence with a non-degenerate, bounded *-representation (whatever this means, but note that this new representation is not unitary) π_ρ of the algebra $C(G)$. Define π_ρ by the Bochner integral

$$\pi_\rho(f) = \int_0^1 f(x)\rho(x)\mathrm{d}x$$

In other words, f acts on v by "convolution against v":

$$f \cdot v = \text{``}f * v\text{''} = \int_0^1 f(x)(x \cdot v)\mathrm{d}x$$

In this setting the action of the character χ_n (which is an element of $C(G)$) on $\mathcal{H}$ is an orthogonal idempotent; that is, the map $\chi_n * -$ satisfies

$$(\chi_m * -)(\chi_n * -) = (\chi_m * \chi_n) * - = \delta_{mn}\chi_n * -$$

To see this, use the Fourier transform: It suffices to see that $\widehat{\chi_m * \chi_n} = \delta_m \cdot \delta_n = \delta_{mn}\delta_n$ (here δ_{mn} is the usual Kronecker delta).

The subspaces $\mathcal{H}_n = \chi_n * \mathcal{H}$ of $\mathcal{H}$ combine to form $\mathcal{H}^{\mathrm{alg}} = \bigoplus_{n\in\mathbb{Z}} \mathcal{H}_n$. It turns out that $\mathcal{H}^{\mathrm{alg}}$ is dense in $\mathcal{H}$. As expected, $\mathrm{U}(1)$ acts by pointwise multiplication by characters on the isotypic components of $\mathcal{H}$, the summands in $\mathcal{H}^{\mathrm{alg}}$. Like before, we can get a sheaf picture where to each point n of $\mathbb{Z}$ we place above it the vector space $\mathcal{H}_n$. The picture is a little different since the Hilbert space $\mathcal{H}$ is the closure of $\mathcal{H}^{\mathrm{alg}}$.

$$\begin{array}{ccccc} \mathcal{H} = & \overline{\mathcal{H}_1 \oplus} & \overline{\mathcal{H}_2 \oplus} & \overline{\cdots \oplus} & \overline{\mathcal{H}_\ell} \\ \mathbb{Z} & 1 & 2 & \cdots & \ell \end{array}$$

The density of $\mathcal{H}^{\mathrm{alg}}$ is due to some version of Plancherel's theorem in this setting. The algebra $C(G)$ does not have a unit, but if it did, it should have been the delta function δ_{1_G} (which is not a continuous function, or much less a function at all). This is evidence that we should have taken some larger space to be the analogue of the group algebra in this setting; that is, perhaps some space of distributions (with convolution as product) would have been a better choice. If we believe that such an analogue exists, then δ_{1_G} acts as the identity operator on $\mathcal{H}$:

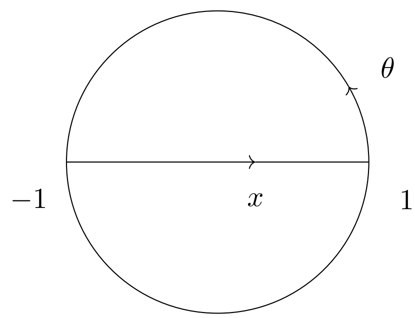
$$\mathrm{id}_{\mathcal{H}}\, v = v = \delta_{1_G} * v$$

But by the Fourier transform δ_{1_G} corresponds to the unit for multiplication in the analogue of the group algebra for $\mathbb{Z}$ with pointwise multiplication, which is the constant function $\mathbf{1}_\mathbb{Z}$ ($\mathbf{1}_\mathbb{Z}(n) = 1$, notably not finitely supported). Since $\mathbf{1}_\mathbb{Z} = \sum_{n\in\mathbb{Z}} \delta_n$, taking the inverse Fourier transform shows that $\mathrm{id}_\mathcal{H} = \sum_{n\in\mathbb{Z}} \chi_n * -$; this shows that $\mathcal{H}^{\mathrm{alg}}$ is dense in $\mathcal{H}$.

Chebyshev polynomials of the first kind

The Chebyshev polynomials (of the first kind) are special functions, like the characters given by $n \mapsto \exp(2\pi i n x)$ or $x \mapsto \exp(2\pi i n x)$ that interact nicely with the Fourier theory.

Comparing the real and imaginary parts in de Moivre's formula $(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$, one can prove that $\cos(n\theta)$ is a polynomial in $\cos(\theta)$. The change of coordinates $x = \cos(\theta) = (\exp(i\theta) + \exp(-i\theta))/2$ can now be used to turn power series in $\exp(i\theta), \exp(-i\theta)$ (Fourier series) representing even functions on the circle to power series in x. We can use a different change of coordinates $\sin(\theta) = (\exp(i\theta) - \exp(-i\theta))/2i$ to handle odd functions on the circle; these will lead to Chebyshev polynomials of the second kind.



The n-th Chebyshev polynomial of the first kind $T_n(x)$ is defined to be the polynomial for which

$$T_n(\cos(\theta)) = \cos(n\theta)$$

That $\{\cos(n\theta)\}$ is an orthogonal basis for a suitable space of even functions on the circle corresponds to $\{T_n(x)\}$ being an orthogonal basis of a suitable space of functions on $[-1, 1]$ with weighted L^2 inner product $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)/\sqrt{1-x^2}\mathrm{d}x$. Some of the first few Chebyshev polynomials of the first kind are

$$\begin{aligned} T_0(x) &= 1 & T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 & T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 & T_5(x) &= 16x^5 - 20x^3 + 5x \end{aligned}$$

We should think of the $T_n(x)$ as characters; in particular we can undo the change of coordinates $x = \cos(\theta)$ to obtain the universal character $(n, \theta) \mapsto T_n(\cos(\theta)) = \cos(n\theta) = (\exp(in\theta) - \exp(-in\theta))/2$.

These polynomials satisfy the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

which should be thought of as a second order difference equation in n; that is, a discrete or otherwise integral version of a second order ODE. The Fourier transform takes shift operators to multiplier operators, so in this setting, undoing the change of coordinates $x = \cos(\theta)$ and applying a suitable Fourier transform to the resulting recurrence relation would return a second order algebraic equation (a quadratic equation) for $\cos(n\theta)$ in θ.

These polynomials also satisfy the differential equation

$$(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}T_n(x) - x\frac{\mathrm{d}}{\mathrm{d}x}T_n + n^2 T_n = 0$$

in x, so we should obtain a second order algebraic equation in n after applying a suitable Fourier transform to the above differential equation. That the above difference equation and differential equation are of second order should not be surprising since the universal character is $\cos(n\theta)$ in this setting.

From the point of view of trigonometric polynomials, for some space of functions the set

$$\{\underline{1}, \underline{\cos(\theta), \sin(\theta)}, \underline{\cos(2\theta), \sin(2\theta)}, \dots\}$$

is an orthogonal basis. The underlined groups of functions are eigenspaces for the Laplacian $\Delta_{\mathrm{U}(1)} = \frac{\partial^2}{\partial\theta^2}$ on the circle. We can obtain these functions in a different way by first considering harmonic polynomials in the plane; that is, polynomials $p(x, y)$ for x, y real which satisfy $\Delta_{\mathbb{R}^2} p = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)p = 0$. Then consider the homogeneous polynomials in the plane; these are the polynomials p for $p(kx, ky) = k^{\deg(p)}p(x, y)$; in polar coordinates these polynomials are separable with $p(r, \theta) = r^{\deg p}\tilde{p}(\theta)$ for some function $\tilde{p}(\theta)$ ($\deg p$ is the homogeneous degree of p). If p is a harmonic homogeneous polynomial,

$$0 = \Delta_{\mathbb{R}} p(r, \theta) = \left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{1}{r^2}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\right) r^{\deg p}\tilde{p}(\theta) = r^{\deg p - 2}\left((\deg p)^2\tilde{p}(\theta) + \frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\tilde{p}(\theta)\right)$$

By restricting to the circle $r = 1$, observe that $\tilde{p}$ is an eigenfunction of the Laplacian on the circle, which by the theory of ordinary differential equations tells us that $\tilde{p}$ is some linear combination of $\cos(n\theta)$ and $\sin(n\theta)$ for $n = \deg p$.

That the real eigenspaces (that aren't $\mathbb{R}\{1\}$) for the Laplacian on the circle are two-dimensional comes from the action of the Laplacian on a suitable function space on the circle. Other differential operators may have different eigenspaces, of possibly different dimensions. In a Lie algebra, there is no notion of squaring elements, only taking their bracket. The universal enveloping algebra for a Lie algebra is an algebra which makes it possible to multiply elements of the Lie algebra together, and from a representation of a Lie group we should obtain a module over the universal enveloping algebra in a similar way to how we did so with the group algebra. We saw before how to obtain the derivative $\frac{\mathrm{d}}{\mathrm{d}\theta}$ from the action of the Lie algebra of $\mathrm{U}(1)$ on functions. In the universal enveloping algebra it makes sense to square the element producing the derivative (it was -1 for the Lie algebra of $\mathbb{R}/\mathbb{Z}$ or $-i$ for the Lie algebra of $\mathrm{U}(1)$), which would then act on functions as the Laplacian.

Lecture 6 September 11

The Fourier transform on $\mathbb{R}$

Let G be the real line $\mathbb{R}$, which henceforth we will denote by $\mathbb{R}_x$ to indicate that the coordinate on $\mathbb{R}$ is x. Like before, because G is an LCA group we know that the irreps of G are its one-dimensional unitary representations, which are parameterized by $\widehat{G} = \mathbb{R}_t$ (again, this notation means the real line with coordinate t, to distinguish this real line with the real line defining G). The universal character χ in this setting is the map $(x,t) \mapsto \exp(ixt)$.

Let $\mathcal{H}$ be a unitary representation of G; that is, $\mathcal{H}$ is a Hilbert space that the real line acts on by unitary operators. In this case, the image of $\mathbb{R}$ in $\mathrm{U}(\mathcal{H})$ is a one-parameter subgroup $\{U_x\}_{x\in\mathbb{R}_x}$ of $\mathrm{U}(\mathcal{H})$. By Stone's theorem on one-parameter unitary groups (Wikipedia), it follows that $U_x = \exp(ixH)$ for some self-adjoint operator H on $\mathcal{H}$. So the data of a unitary representation $\mathcal{H}$ of G corresponds to the data of a single self-adjoint operator on $\mathcal{H}$. By differentiating this action, that is, passing to the action of the Lie algebra of G on $\mathcal{H}$, we find that the Lie algebra of G acts by the operator iH on $\mathcal{H}$; this element (in physics, just H) is called the infinitesimal generator for the group $\{U_x\}_{x\in\mathbb{R}_x}$.

One example of this story is in quantum mechanics, where $\mathcal{H}$ is the space of states for a quantum mechanical system and $\mathbb{R}$ acts on $\mathcal{H}$ by time evolution. The infinitesimal generator of the one-parameter group $\{U_x\}_{x\in\mathbb{R}_x}$ in this setting is the system's Hamiltonian.

From the spectral theorem for self-adjoint operators on Hilbert spaces, we obtain a kind of dictionary between

$$\{\text{irreps of unitary reps}\} \longleftrightarrow \{\text{eigenvalues in } \mathbb{R}\}$$

where the left side lives in the world of representation theory and the right side lives in the world of self-adjoint operator theory. Specifically, for a fixed unitary representation $\mathcal{H}$ of G, decomposing $\mathcal{H}$ into its irreps amounts to finding the eigenvalues of the operator H defining the representation. More generally, the decomposition of unitary representations $\mathcal{H}$ of G corresponds to the spectrum of of the operator H defining the representation. We will see this later.

The universal character χ is used to obtain the Fourier transform $\mathrm{Fun}(G) \to \mathrm{Fun}(\widehat{G})$ as usual. It is given by $f \mapsto \hat{f}$ where

$$\hat{f}(t) = \int_{\mathbb{R}_x} f(x)\exp(itx)\mathrm{d}x,$$

and the Fourier inversion theorem in this setting is

$$f(x) \approx \int_{\mathbb{R}_t} \hat{f}(t)\exp(-itx)\mathrm{d}t,$$

where as usual we should worry about convergence. We should also care about what functions we apply the Fourier transform to. For example, a nice result from analysis is that the Fourier transform defines a unitary isomorphism of $L^2(\mathbb{R}_x)$ with $L^2(\mathbb{R}_t)$ (that the isomorphism is unitary is usually known as Plancherel's theorem).

The Fourier transform diagonalizes the translation operators like before. The group G acts naturally on $\mathrm{Fun}(G)$ by translation:

$$y \cdot f =: \tau_y f = f((-) - y) \quad \text{for } y \in \mathbb{R}_x$$

and one can calculate the infinitesimal version of this as differentiation. The group algebra version of this action is given by convolution; that is, an element f in the group algebra acts on $\mathrm{Fun}(G)$ by the operator $f * -$. By group algebra we might mean a nicer function space like continuous or smooth functions with compact support or even L^1 functions, so it makes sense to convolve, but these are matters of analysis. The point is that the Fourier transform takes these translation actions and simultaneously diagonalizes them, turning them into multiplication operators on $\mathrm{Fun}(\widehat{G})$. In particular, translation τ_y is sent to pointwise multiplication by χ_y, differentiation is sent to multiplication by $-it$ (meaning pointwise multiplication by the map $-i(t \mapsto t)$), and convolution against f is sent to pointwise multiplication by $\hat{f}$.

The statement that $L^2(\mathbb{R}_x)$ is isomorphic to $L^2(\mathbb{R}_t)$ can instead be thought of as a consequence of the spectral theory of the operator $H = i\frac{\mathrm{d}}{\mathrm{d}x}$. This operator should be thought of as an infinitesimal translation operator, which under the Fourier transform is sent to multiplication by t (a diagonal operator). The exponential of iH, $\exp\left(-y\frac{\mathrm{d}}{\mathrm{d}x}\right)$, is the translation operator τ_y; the exponential of it, $\exp(iyt)$, is the multiplication operator χ_y.

An algebraic version of spectral theory

Let V be a complex vector space, and fix an element $T \in \mathrm{End}(V)$. Then regard V as a $\mathbb{C}[t]$-module, where $tv = Tv$ (this formalism is the prototype for the functional calculus). The action of $\mathbb{C}[t]$ on V factors through $\mathbb{C}[t]/\operatorname{Ann}_{\mathbb{C}[t]}(V)$, but this ring is isomorphic to $\mathbb{C}[\mathrm{Spec}(T)]$, the ring of regular functions on the spectrum of the operator T, a subset of $\mathbb{C}$. The isomorphism is to take a polynomial $f(t)$ and send it to the function f as a function of $t \in \mathbb{C}$. These regular functions remember the multiplicities of elements of $\mathrm{Spec}(T)$. If the factor $(t-\lambda)^n$ appears in $m(t) \in \operatorname{Ann}_{\mathbb{C}[t]}(V)$, then the polynomials $f(t), (t-\lambda)f(t), \ldots, (t-\lambda)^{n-1}f(t)$ do not vanish at $t = \lambda$ when viewed as functions on $\mathbb{C}$. In other words, f may have poles (so maybe these are not called regular functions).

The action of f on V will be more clear later, but here is part of it: We have $f \cdot v = f(\lambda)v$ for $v \in V_\lambda$ and $\lambda \in \mathrm{Spec}(T)$, where V_λ is the λ-isotypic component of V. The decomposition of V in this setting is nicest if T is diagonalizable.

Suppose now that V is finitely generated as a $\mathbb{C}[t]$-module. Since $\mathbb{C}[t]$ is a principal ideal domain, we can use the structure theorem for finitely generated modules over principal ideals domains to obtain a "spectral decomposition" of V:

$$V \cong V^{\text{free}} \oplus V^{\text{torsion}} \quad (\text{as } \mathbb{C}[t]\text{-modules})$$

The free part of V corresponds to the "continuous" spectrum of T and the torsion part of V corresponds to the "discrete" spectrum of T. Note that so far we have not thought about whether T was self-adjoint or not; this is all happening purely algebraically. We will not pursue further comparisons with the continuous or point spectrum of bounded self-adjoint operators on Hilbert spaces. Assume now that V is finite-dimensional so that V^{free} in the decomposition above is trivial ($\mathbb{C}[t]$ is infinite-dimensional as a complex vector space). Then from the same structure theorem for finitely generated modules over principal ideals we have

$$V \cong \bigoplus_{i=1}^{m} \frac{\mathbb{C}[t]}{(t-\lambda_i)^{n_i}},$$

where the λ_i may repeat. This decomposition of V is called the primary decomposition of V, from which we can obtain the Jordan normal form of T. The Jordan normal form of T is a block matrix

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{pmatrix}$$

where each J_i is the $n_i \times n_i$ matrix

$$\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

If $n_i > 1$, we say that λ_i is a generalized eigenvalue of T.

For example, the natural action of t on the $\mathbb{C}[t]$-module $\mathbb{C}[t]/(t-\lambda)^n$ with $\mathbb{C}$-basis $\{1, (t-\lambda), \ldots, (t-\lambda)^{n-1}\}$ as a linear operator is given by the $n \times n$ Jordan block with diagonal entries λ.

An analyst's point of view on this spectral decomposition is that the module $\mathbb{C}[t]/(t-\lambda)^n$ is isomorphic to the $\mathbb{C}[t]$-module generated by the $(n-1)$-th distributional derivative of the delta distribution concentrated at λ. The action of $f \in \mathbb{C}[t]$ on a distribution ϕ is given by $(f\phi)(\varphi) = \phi(f\varphi)$, which is usually written in analysis as $\langle f\phi, \varphi\rangle = \langle \phi, f\varphi\rangle$.

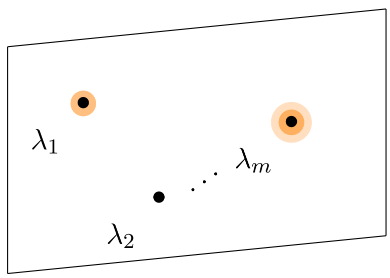
The defining property of the delta distribution δ_λ, which evaluates functions at λ, is that it is annihilated by the ideal $(t-\lambda)$; that is, $(t-\lambda)\delta_\lambda$ is the zero distribution. In other words, these delta distributions are eigenvectors for the action of t. By the n-th distributional derivative of δ_λ, denoted $\delta_\lambda^{(n)}$, we mean the distribution that acts on test functions by the formula $\delta_\lambda^{(n)}(\varphi) = (-1)^n \delta_\lambda(\varphi^{(n)}) = (-1)^n \varphi^{(n)}(\lambda)$. The calculations

$$\begin{aligned} t\delta_\lambda(\varphi) &= (t\varphi)(\lambda) = \lambda\varphi(\lambda) = \lambda\delta_\lambda \\ t\delta'_\lambda(\varphi) &= -(t\varphi)'(\lambda) = -\lambda\varphi'(\lambda) - \varphi(\lambda) = \lambda\delta'_\lambda(\varphi) - \delta_\lambda(\varphi) \\ t\delta''_\lambda(\varphi) &= (t\varphi)''(\lambda) = \lambda\varphi''(\lambda) + 2\varphi'(\lambda) = \lambda\delta''_\lambda(\varphi) - 2\delta'_\lambda(\varphi) \\ &\vdots \end{aligned}$$

show that derivatives of delta distributions are generalized eigenvectors for the action of t; that is, $(t-\lambda)^k\delta_\lambda^{(k-1)} = 0$. As a result the action of t on the complex vector space spanned by $\{\delta'_\lambda, -\delta_\lambda\}$ is given by the matrix $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$. Thus for the case $n = 1$, we have an isomorphism of $\mathbb{C}[t]$-modules $\mathbb{C}[t]/(t-\lambda)^2 \to \mathbb{C}[t]\delta'_\lambda$ given by $1 \mapsto \delta'_\lambda$. In general, t

acts on the complex vector space spanned by $\{(-1)^n\delta_\lambda^{(n-1)}/(n-1)!, \dots, \delta'_\lambda, -\delta_\lambda\}$ by the $n \times n$ Jordan block with diagonal entries λ. Like before, we obtain an isomorphism of $\mathbb{C}[t]$-modules $\mathbb{C}[t]/(t-\lambda)^n \to \mathbb{C}[t]\delta_\lambda^{(n-1)}$ given by $1 \mapsto \delta_\lambda^{(n-1)}$.

So the discrete spectrum of an operator should correspond to derivatives of delta distributions supported at generalized eigenvalues of T. But even though these distributions are supported at points, their values on a test function φ depend on the values of the derivatives of φ (in general distributions may depend on the Taylor series for φ), so the delta distributions are "localized" in this sense. This story can be retold in the language of sheaf theory by thinking of the delta distributions as a sheaf supported at the generalized eigenvalues of T; that is, on $\operatorname{Spec}(\mathbb{C}[t]/\operatorname{Ann}_{\mathbb{C}[t]}(V))$ (here Spec takes the prime spectrum of a ring). The higher the order of the generalized eigenvalue, the higher its "multiplicity" is in this setting (and in the picture below, the multiplicity is indicated by concentric disks).



The infinitesimal action of $G = \mathbb{R}_x$ on $\operatorname{Fun}(\widehat{G})$ by pointwise multiplication by t comes from the Fourier transform of the infinitesimal action of G on $\operatorname{Fun}(G)$ by differentiation. The eigenfunctions of these actions are also expected to be interchanged via the Fourier transform. The delta functions δ_λ for $\lambda \in \mathbb{R}_t$ are eigenfunctions for pointwise multiplication by t; and on the other side of the Fourier transform, the exponentials $\exp(-i\lambda x)$ for $\lambda \in \mathbb{R}_x$ are eigenfunctions of $i\frac{\mathrm{d}}{\mathrm{d}x}$. Some analysis is needed to make this rigorous, but thinking of $\exp(-i\lambda x)$ as a regular distribution (i.e. as an element of a suitably chosen candidate for the group algebra) we have

$$\langle \widehat{\exp(-i\lambda(-))}, \varphi\rangle = \langle \exp(-i\lambda(-)), \hat{\varphi}\rangle = \int_{\mathbb{R}_x} \hat{\varphi}(x)\exp(-i\lambda x)\mathrm{d}x = \varphi(\lambda) = \langle \delta_\lambda, \varphi\rangle$$

so indeed the Fourier transform of $\exp(-i\lambda(-))$ is δ_λ.

Since $G = \mathbb{R}_x$ is not compact, indecomposable representations of G need not coincide with the irreps of G. One way to obtain these is by taking the inverse Fourier transform of derivatives of delta distributions. Some calculations show that the Fourier transform of the function given by $(ix)^k \exp(-i\lambda x)$ is $\delta_\lambda^{(k)}$. By taking the inverse Fourier transform of the calculations from before involving $t\delta_\lambda^{(k)}$, we obtain the equations

$$\begin{aligned} i\frac{\mathrm{d}}{\mathrm{d}x}[\exp(-i\lambda x)] &= \lambda \exp(-i\lambda x) \\ i\frac{\mathrm{d}}{\mathrm{d}x}[(ix)\exp(-i\lambda x)] &= \lambda(ix)\exp(-i\lambda x) - \exp(-i\lambda x) \\ i\frac{\mathrm{d}}{\mathrm{d}x}[(ix)^2\exp(-i\lambda x)] &= \lambda(ix)^2\exp(-i\lambda x) - 2(ix)\exp(-i\lambda x) \\ &\vdots \end{aligned}$$

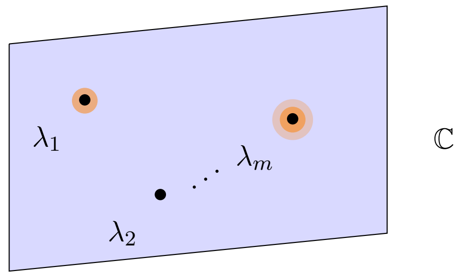
from which we can see that $i\frac{\mathrm{d}}{\mathrm{d}x}$ acts on the complex vector space spanned by $\{(-1)^n(ix)^{n-1}\exp(-i\lambda x)/(n-1)!, \ldots, (ix)\exp(-i\lambda x), -\exp(-i\lambda x)\}$ by the $n \times n$ Jordan block with diagonal entries λ. Similarly, if we think of this vector space instead as a $\mathbb{C}[X]$-module where X acts by $i\frac{\mathrm{d}}{\mathrm{d}x}$, there is an isomorphism $\mathbb{C}[X]/(X-\lambda)^n \to \mathbb{C}[X](ix)^{n-1}\exp(-i\lambda x)$ given by $1 \mapsto (ix)^{n-1}\exp(-i\lambda x)$. Notice this is all completely symmetric to what we did earlier with the delta distributions. The indecomposable representation we have obtained is the n-dimensional representation of the Lie algebra of $\mathbb{R}_x$ given by the $\mathbb{C}$-span of $\{(ix)^{n-1}\exp(-i\lambda x), \ldots, (ix)\exp(-i\lambda x), \exp(-i\lambda x)\}$ (this is the induced Lie algebra representation given by acting by scalar multiples of $\frac{\mathrm{d}}{\mathrm{d}x}$). It is the case that this vector space is also an indecomposable representation of the original group $\mathbb{R}_x$ under translation.

So in the usual sheaf picture, given a finite-dimensional representation V of $\mathbb{R}_x$, we can decompose V into the direct sum of indecomposable subrepresentations isomorphic to the ones in the previous paragraph (of possibly different dimensions of course), and each subspace should be viewed as living over the corresponding generalized eigenvalue of the operator H defining the action of $\mathbb{R}_x$ on V.

One objection to the arguments from before is that none of these functions (the exponential, the delta distribution) belong to $L^2(\mathbb{R}_x)$ or are even functions. So these representations above are not necessarily unitary representations; somehow the Fourier theory is able to see non-unitary representations...

To address the continuous spectrum, we return to the setting where V is a finitely generated $\mathbb{C}[t]$-module (where t acts by an operator $T \in \operatorname{End}(V)$), so that V decomposes as $V = V^{\text{free}} \oplus V^{\text{torsion}}$. We saw earlier that the V^{torsion} summand corresponds to the discrete spectrum of the operator T, so on the other side the V^{free} summand corresponds to the continuous spectrum of T.

Since V is finitely generated as a $\mathbb{C}[t]$-module, $V^{\text{free}} \cong \mathbb{C}[t]^\ell$ for ℓ finite. Note that $\operatorname{Spec}(\mathbb{C}[t]/(t-\lambda)^n)$ is the point $\{\lambda\}$ (thought of as having multiplicity n), and that $\operatorname{Spec}(\mathbb{C}[t])$ is the affine line $\mathbb{C}$. So in contrast to the discrete spectrum, the continuous spectrum in this setting "lives everywhere" over $\mathbb{C}$ (in blue).



There are no eigenvectors under the action of t in the free module $\mathbb{C}[t]$; that is, there are no annihilators of V^{free} under the action of T. However, there will be "co-eigenvectors" or eigenfunctionals belonging to $\mathbb{C}[t]^* = \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}[t], \mathbb{C})$; for example, the delta distribution δ_λ. This is to say there are no subspaces of $\mathbb{C}[t]$ that are eigenspaces under the action of t, but there are quotients of $\mathbb{C}[t]$ that are eigenspaces: The functional δ_λ defines a surjection $\mathbb{C}[t] \xrightarrow{\delta_\lambda} \mathbb{C}$, and by the first isomorphism theorem we obtain the quotient eigenspace $\mathbb{C}[t]/(t-\lambda)$. Indeed, $t = (t-\lambda)+\lambda$ acts on the $\mathbb{C}[t]$-module $\mathbb{C}[t]/(t-\lambda)$ by scalar multiplication by λ. So in some sense this sheaf that "lives everywhere" is some kind of direct integral of delta sheaves living at every point.

Here are some examples that illustrate why V should be finitely generated as a $\mathbb{C}[t]$-module. Think of the ring $\mathbb{C}[t]$ as an $\mathbb{N}$-graded ring consisting of $\mathbb{C}$ in each degree or graded component, where t acts by moving elements of $\mathbb{C}$ up one degree:

$$\begin{array}{ccccccccc} 0 & & 1 & & 2 & & 3 & & \\ \mathbb{C} & \xrightarrow{t\cdot} & \mathbb{C} & \xrightarrow{t\cdot} & \mathbb{C} & \xrightarrow{t\cdot} & \mathbb{C} & \xrightarrow{t\cdot} & \cdots \end{array}$$

In contrast the ring $\mathbb{C}[t,t^{-1}]$ can be thought of a $\mathbb{Z}$-graded ring consisting of $\mathbb{C}$ in each degree, where t acts by moving elements of $\mathbb{C}$ up one degree, and t^{-1} acts by moving elements of $\mathbb{C}$ down one degree:

$$\begin{array}{ccccccccccc} & & -2 & & -1 & & 0 & & 1 & & \\ \cdots & \underset{t^{-1}\cdot}{\overset{t\cdot}{\rightleftarrows}} & \mathbb{C} & \underset{t^{-1}\cdot}{\overset{t\cdot}{\rightleftarrows}} & \mathbb{C} & \underset{t^{-1}\cdot}{\overset{t\cdot}{\rightleftarrows}} & \mathbb{C} & \underset{t^{-1}\cdot}{\overset{t\cdot}{\rightleftarrows}} & \mathbb{C} & \underset{t^{-1}\cdot}{\overset{t\cdot}{\rightleftarrows}} & \cdots \end{array}$$

As a $\mathbb{C}[t]$-module, $\mathbb{C}[t,t^{-1}]$ is not finitely generated due to the uncountably many copies of $\mathbb{C}$ in negative degrees. Further, there is no $\mathbb{C}[t]$-module quotient of $\mathbb{C}[t,t^{-1}]$ that is concentrated exactly in degree zero (that is, no eigenquotients under the action of t); that is, there is no evaluation map $\mathbb{C}[t,t^{-1}] \to \mathbb{C}$ that sends t to zero. The algebraic dual space of $\mathbb{C}[t,t^{-1}]$ is the ring of formal power series over $\mathbb{Z}$, denoted $\mathbb{C}[[t,t^{-1}]]$. The delta functional δ_λ lives in this ring for $\lambda \neq 0$, so we obtain a sheaf that lives everywhere except at the origin; we know this because $\mathrm{Spec}(\mathbb{C}[t,t^{-1}])$ is $\mathbb{C}^\times$ (since $\mathbb{C}[t,t^{-1}]$ is the localization of $\mathbb{C}[t]$ at the ideal (t)).

Another example is the ring $\mathbb{C}(t)$, which is not finitely generated as a $\mathbb{C}[t]$-module. There are no $\mathbb{C}[t]$-submodules or quotients of $\mathbb{C}(t)$ concentrated in degree 0 (no eigenspaces or eigenquotients under the action of t). The spectrum of $\mathbb{C}(t)$ is the origin, because it is a field; or observe that we have inverted too many elements in $\mathbb{C}[t]$ (localized away from the zero ideal), so the sheaf we obtain in this case lives over the origin. By assuming V is finitely generated as a $\mathbb{C}[t]$-module, we are led to study a quasicoherent sheaf over $\operatorname{Spec}(\mathbb{C}[t]/\operatorname{Ann}_{\mathbb{C}[t]}(V))$ whose global sections are V.

An analytic version of spectral theory

Let $\mathcal{H}$ be a Hilbert space. The spectral theorem for self-adjoint operators on Hilbert spaces gives a correspondence

$$\{\text{self-adjoint operators } H \text{ on } \mathcal{H}\} \longleftrightarrow \{\text{measurable fields of Hilbert spaces } \{\mathcal{H}_t \mid t \in \mathbb{R}\} \text{ over } \mathbb{R}\}$$

A measurable field over $\mathbb{R}$ is a family of Hilbert spaces $\mathcal{H}_t$ for each $t \in \mathbb{R}$ together with a subspace M of $\prod_{t\in\mathbb{R}} \mathcal{H}_t$ which denote the measurable sections (the elements of m should be thought of as functions), which satisfy a few properties involving measurability (see the nLab). In some sense the assignment $t \mapsto \mathcal{H}_t$ should be measurable and $\mathcal{H} = \int_{\mathbb{R}}^{\oplus} \mathcal{H}_t \mathrm{d}t$, a direct integral (a continuous analogue of a direct sum).

One idea in functional calculus is to make a commutative algebra out of an operator; this is analogous to the idea of a group algebra. This leads to an action of the von Neumann algebra $L^\infty(\mathbb{R})$ on $\mathcal{H}$ by integrating functions against a projection-valued measure and applying the resulting projection; in this way $\mathcal{H}$ becomes a W^*-module.

A projection-valued measure on a measure space X is a function that takes in measurable sets E of X and returns orthogonal projectors π_E onto certain subspaces of a Hilbert space $\mathcal{H}$. This function needs to take the empty set to the zero operator, X to the identity map, disjoint unions to sums of projectors, and finite intersections to products of projections (compositions of projections; which commute). If H is a self-adjoint operator on $\mathcal{H}$, then there is a unique projection-valued measure π_H on $\mathbb{R}$ for which $H = \int_{\mathbb{R}} \lambda \mathrm{d}\pi_H(\lambda)$ (and this integral is actually just taken over the spectrum of H, which is compact).

It turns out that the map $f \mapsto \int_{\mathbb{R}} f(\lambda)\mathrm{d}\pi_H(\lambda)$ from $L^\infty(\mathbb{R})$ to bounded operators on $\mathcal{H}$ is a homomorphism of C^*-algebras, so in particular the product of functions goes to the composition of the resulting operators (and the composition commutes; we should think of this map as returning diagonal operators).

Returning to our unitary representation $\mathcal{H}$, the action of $\mathbb{R}_x$ on $\mathcal{H}$ is determined by the self-adjoint operator H for which $x \in \mathbb{R}_x$ acts by $U_x = \exp(ixH)$. Then such an H gives us a unique projection-valued measure π_H which we use to define the action of $L^\infty(\mathbb{R})$ on $\mathcal{H}$.

For example, if I is an interval in $\mathbb{R}$, then the characteristic function $\chi_I \in L^\infty(\mathbb{R})$ acts on $\mathcal{H}$ by the operator $\int_I \mathrm{d}\pi_H(\lambda)$, which is an idempotent since $\chi_I^2 = \chi_I$ and integration against π_H preserves products. Compare this with the operator given by convolution against a character in the case when $G = \mathrm{U}(1)$. Think of the operator $\int_I \mathrm{d}\pi_H(\lambda)$ as projecting onto the piece of $\mathcal{H}$ that lives over the interval I.

Compare this with the case where $\mathcal{H}$ is finite-dimensional and H is diagonalizable, in which case the direct integral is reduced to the direct sum and the projection we get by integrating the characteristic function of an eigenvalue λ against the projection-valued measure π_H is the projection to the eigenspace $\mathcal{H}_\lambda$. If we like we can integrate other characteristic functions to get direct sums of eigenspaces.

In general $\mathcal{H}$ need not have eigenspaces or eigenquotients, and the spectral theory of the operator H coming from the group action will tell us how $\mathcal{H}$ is smeared over the spectrum of H, and by using π_H we can obtain projections to parts of $\mathcal{H}$ living over the continuous spectrum. Notice that if t belongs to the continuous spectrum of H, we cannot obtain a projection onto $\mathcal{H}_t$ since such an operator would have to come from integrating a function concentrated at $\lambda = t$ against π_H, which yields the zero operator since the measure of a point is zero.

The way $\mathcal{H}$ is smeared over the spectrum of H can be thought of as an analyst's version of a sheaf. We can think of the atoms of the sheaf as eigenspaces (from delta functions), and the molecules as the spaces coming from generalized eigenvalues (from derivatives of delta functions). Unitary representations need not have either of these in the case the discrete spectrum of the operator H defining the representation is empty, but in general the representation is smeared over the continuous spectrum of H.

A tiny preview of the Heisenberg group

The group $\mathbb{R}_x$ acts on $L^2(\mathbb{R}_x)$ by translations $\tau_y = \delta_y * -$ for $y \in \mathbb{R}_x$, and on the other side of the Fourier transform, $\mathbb{R}_x$ acts on $L^2(\mathbb{R}_t)$ by pointwise multiplication by characters $\chi_y = \chi(y, -)$. Due to Pontryagin duality we can also think about the action of the dual group $\mathbb{R}_t$ on $L^2(\mathbb{R}_t)$ by translation and on the other side of the Fourier transform on $\mathbb{R}_t$ (not the inverse Fourier transform!), $\mathbb{R}_t$ acts on $L^2(\mathbb{R}_x)$ by pointwise multiplication by characters $\chi_t = \chi(-, t)$ for $t \in \mathbb{R}_t$.

These observations may be collected on the side $G = \mathbb{R}_x$ to see that the group $G \times \widehat{G}$ acts on $L^2(G)$. One natural question is to ask if the action of G commutes with the action of $\widehat{G}$ on $L^2(G)$ (where $G, \widehat{G}$ are henceforth viewed as subgroups of $G \times \widehat{G}$). The actions of $G, \widehat{G}$ on $L^2(G)$ should be thought of as belonging to the multiplicative group $\mathrm{Aut}(\mathrm{Fun}(G))$, so to ask if they commute is to calculate the commutator of the action of group elements $g \in G$ and $\hat{g} \in \widehat{G}$ denoted by $[g\cdot, \hat{g}\cdot] = [g, \hat{g}]\cdot$. On a test function, we have

$$\begin{aligned}
[g, \hat{g}] \cdot f &= g^{-1}\hat{g}^{-1} g\hat{g} \cdot f \\
&= g^{-1}\hat{g}^{-1} g(\chi(-, \hat{g}) f) \\
&= g^{-1}\hat{g}^{-1}(\overline{\chi(g, \hat{g})}\chi(-, \hat{g}) f(g^{-1}-)) \\
&= g^{-1}(\chi(-, \hat{g}^{-1})\overline{\chi(g, \hat{g})}\chi(-, \hat{g}) f(g^{-1}-)) \\
&= \overline{\chi(g, \hat{g})} f(gg^{-1}-) \\
&= \overline{\chi(g, \hat{g})} f
\end{aligned}$$

So the action of G and $\widehat{G}$ differ by multiplication by $\overline{\chi(g, \hat{g})} \in \mathrm{U}(1)$. To account for these additional symmetries, we obtain the Heisenberg group via a central extension of $G \times \widehat{G}$ by $\mathrm{U}(1)$ (or if G is instead a finite Abelian group, we only instead use the roots of unity that $\overline{\chi(g, \hat{g})}$ lie in). The central extension that yields the Heisenberg group is specifically the short exact sequence of groups

$$1 \to \mathrm{U}(1) \to \mathrm{Heis} \to G \times \widehat{G} \to 1$$

Note that U(1) is central in Heis, that is, viewing U(1) as a subgroup of Heis, the elements in U(1) belong to the center $Z(\text{Heis})$ of the Heisenberg group (since the action of multiplication by scalars commuted with the actions of $G, \widehat{G}$ to begin with).

The Heisenberg group acts on $L^2(G)$ in three ways: by scalar multiplication, by translation, and by multiplication by characters (coming from $\mathrm{U}(1), G, \widehat{G}$ respectively). One advantage of forming the Heisenberg group this way is that it will be easy to show that $L^2(G)$ is actually an irreducible representation of the Heisenberg group, and many results surrounding harmonic analysis are much easier to prove.

The discussion above about the commutator of the group actions of $G, \widehat{G}$ can be differentiated to recover a fundamental commutation relation the physicists and analysts are deeply familiar with. Let $\mathfrak{g}, \widehat{\mathfrak{g}}$ denote the Lie algebras of $G, \widehat{G}$ respectively. The infinitesimal actions of $\mathfrak{g}, \widehat{\mathfrak{g}}$ on $\mathrm{Fun}(G)$ are given by the following: $y \in \mathfrak{g}$ acts on $f \in \mathrm{Fun}(G)$ by $y \cdot f = \frac{\mathrm{d}}{\mathrm{d}t}(\exp(yt) \cdot f)|_{t=0} = Df(-y)$ and $\hat{y} \in \widehat{\mathfrak{g}}$ acts on $f \in \mathrm{Fun}(G)$ by $\hat{y} \cdot f = \frac{\mathrm{d}}{\mathrm{d}t}(\exp(\hat{y}t) \cdot f)|_{t=0} = \big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})f$. View $\mathfrak{g}, \widehat{\mathfrak{g}}$ as subspaces of the Lie algebra $\mathfrak{g} \oplus \widehat{\mathfrak{g}}$.

In general, a representation of a Lie group G given by $G \xrightarrow{\rho} \mathrm{GL}(V)$ gives rise to a Lie algebra representation $\mathfrak{g} \xrightarrow{\mathrm{d}\rho} \mathfrak{gl}(V) = \mathrm{End}(V)$ by the above formulas. What is meant by the commutation relation of the actions of $y, \hat{y}$, we mean the commutation relation as operators in $\mathrm{End}(V)$, so we calculate the ordinary commutator of operators. On a test function, we have

$$\begin{aligned}
[y\cdot, \hat{y}\cdot]f &= y \cdot (\hat{y} \cdot f) - \hat{y} \cdot (y \cdot f) \\
&= D_g\big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})(-y)f + \big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})Df(-y) - \big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})Df(-y) \\
&= D_g\big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})(-y)f
\end{aligned}$$

So the commutation relation is $[y\cdot, \hat{y}\cdot] = D_g\big(D_{\hat{g}}\chi(-, 1_{\widehat{G}})\big)(\hat{y})(-y)$ (the right hand side is a number, so we mean scalar multiplication by this number). In the case that $G = \mathbb{R}_x$ and $\widehat{G} = \mathbb{R}_t$, $\chi(x, t) = \exp(ixt)$ so by choosing $y = -1$, $\hat{y} = -i$ we obtain the anticipated commutation relation

$$[(-1)\cdot, (-i)\cdot] = \left[\frac{\mathrm{d}}{\mathrm{d}x}\cdot, x\cdot\right] = 1$$

Lecture 7 September 16

Central extensions

We slowly transition away from Abelian groups by first discussing a slightly non-Abelian group, the Heisenberg group. It is only "slightly" non-Abelian because it is formed as a central extension of two Abelian groups.

In general, a central extension of the group H (which need not be Abelian) by an Abelian group A is a group $\tilde{H}$ which fits in the short exact sequence of groups

$$1 \to A \to \widetilde{H} \to H \to 1$$

(so $A \to \widetilde{H}$ is injective and $\widetilde{H} \to H$ is surjective) with $A \hookrightarrow Z(\widetilde{H})$. If H is Abelian then the group $\widetilde{H}$ is close to being Abelian in the sense that it is a nilpotent group with nilpotency class 2 (Abelian groups have nilpotency class 1).

A nilpotent group G of nilpotency class n is a group whose central series of shortest length terminates in the whole group after n many steps; that is, there is a series of normal subgroups G_i

$$1 = G_0 \lhd G_1 \lhd \cdots \lhd G_n = G$$

with $G_{i+1}/G_i \leq Z(G/G_i)$ (or equivalently $[G, G_{i+1}] \leq G_i$). With this definition it is clear that if H is Abelian, then $\widetilde{H}$ is of nilpotency class 2 since its upper central series is

$$1 \lhd A \lhd \widetilde{H}$$

To classify central extensions of H (which need not be Abelian) by A, we turn to group cohomology. By some process in group cohomology, we can produce a set (really a group) denoted $H^2(H, A)$, which classifies central extensions of H by A up to isomorphism. As a set, $\widetilde{H} \cong H \times A$. The multiplication in $\widetilde{H}$ is given by

$$(g, a)(h, b) = (gh, f(g, h)ab)$$

where $f \colon H \times H \to A$ is a set map satisfying the 2-cocycle conditions

$$f(1_H, 1_H) = 1_A \quad \text{and} \quad f(g, h)f(gh, j) = f(g, hj)f(h, j)$$

for $g, h, j \in H$. The first condition is needed if we want the restriction of the product to A viewed as a subgroup of $\widetilde{H}$, as $\{1_H\} \times A \subseteq \widetilde{H}$, to agree with the product in A. The second condition is to ensure the multiplication is associative. (See chapter 3 of the textbook *A Mathematical Introduction to Conformal Field Theory.*)

It is important to note that a central extension is not necessarily a semidirect product of H and A; the semidirect products of H and A are the split extensions; that is, the extensions

$$1 \to A \to \widetilde{H} \xrightarrow{p} H \to 1$$

where we do not assume $A \hookrightarrow Z(\widetilde{H})$ (or that A is Abelian; i.e., a possibly non-central extension) but there is a homomorphism $s\colon H \to \widetilde{H}$ that is a section; that is, $ps = \mathrm{id}_H$. A central extension of H by A that is also a split extension makes $\widetilde{H}$ isomorphic to the direct product of H and A; that is, $\widetilde{H} \cong H \times A$ (see this MSE question). The split central extensions correspond to the identity element in the group $H^2(H, A)$.

The Heisenberg group

We return to the Heisenberg group. Let G be an LCA group and let $\widehat{G} = \mathrm{Hom}_{\mathbf{Group}}(G, \mathrm{U}(1))$. Then the universal character $\chi(-,-)$ is a map from $G \times \widehat{G}$ to $\mathrm{U}(1)$, and if G is finite the universal character really maps into a finite set of roots of unity μ_n for some n. The Heisenberg group Heis is obtained as a central extension

$$1 \to \mathrm{U}(1) \to \mathrm{Heis} \to G \times \widehat{G} \to 1$$

(where if G is finite, we could replace $\mathrm{U}(1)$ by μ_n if desired) and the 2-cocycle $f\colon (G \times \widehat{G}) \times (G \times \widehat{G}) \to \mathrm{U}(1)$ appearing in the multiplication formula

$$(g, \hat{g}, z)(h, \hat{h}, w) = (gh, \hat{g}\hat{h}, f(g, \hat{g}, h, \hat{h})zw)$$

in Heis (viewed as the set $(G \times \widehat{G}) \times \mathrm{U}(1)$) is given by

$$f(g, \hat{g}, h, \hat{h}) = \overline{\chi(g, \hat{h})}$$

By comparing the results of the following calculations it follows that f really is a 2-cocycle:

$$\begin{aligned} f(g, \hat{g}, h, \hat{h}) f(gh, \hat{g}\hat{h}, j, \hat{j}) &= \overline{\chi(g, \hat{h})\chi(gh, \hat{j})} \\ f(g, \hat{g}, hj, \hat{h}\hat{j}) f(h, \hat{h}, j, \hat{j}) &= \overline{\chi(g, \hat{h}\hat{j})\chi(h, \hat{j})} \end{aligned}$$

We can obtain an isomorphic group by using the 2-cocycle $\tilde{f}$ given by $\tilde{f}(g, \hat{g}, h, \hat{h}) = \chi(h, \hat{g})$ instead of f in the multiplication formula above. Results from group cohomology imply that we obtain isomorphic groups if f and $\tilde{f}$ differ by a 2-coboundary, which is the case:

$$\tilde{f}(g, \hat{g}, h, \hat{h})^{-1} f(g, \hat{g}, h, \hat{h}) = \chi(h, \hat{h})\overline{\chi(gh, \hat{g}\hat{h})}\chi(g, \hat{g})$$

The 2-coboundary in this situation is the image of $\chi\colon G \times \widehat{G} \to \mathrm{U}(1)$ under a particular differential d^1 coming from a bar complex, that is, the 2-coboundary is $d^1\chi\colon (G \times \widehat{G}) \times (G \times \widehat{G}) \to \mathrm{U}(1)$ given by $d^1\chi(g, \hat{g}, h, \hat{h}) = \chi(h, \hat{h})\chi(gh, \hat{g}\hat{h})^{-1}\chi(g, \hat{g})$.

As an aside, some Abelian subgroups of Heis are $G \times \{1_{\widehat{G}}\} \times \mathrm{U}(1)$, $\{1_G\} \times \widehat{G} \times \mathrm{U}(1)$, $G \cong G \times \{1_{\widehat{G}}\} \times \{1_{\mathrm{U}(1)}\}$, and $\widehat{G} \cong \{1_G\} \times \widehat{G} \times \{1_{\mathrm{U}(1)}\}$. The subgroups $G \times \{1_{\widehat{G}}\} \times \mathrm{U}(1)$ and $\{1_G\} \times \widehat{G} \times \mathrm{U}(1)$ do not commute with each other!

In the past, we had three group actions on $\mathrm{Fun}(G)$. One was the action of G by translation, another was the action of $\widehat{G}$ by multiplication by characters, and lastly the usual scalar multiplication by $\mathrm{U}(1)$. The action of $\mathrm{U}(1)$

commutes with the first two group actions (and of course with its own group action). We saw before that the actions of $G, \widehat{G}$ do not commute, which was why we formed the Heisenberg group in the first place. These three group actions can be thought of as living in $\mathrm{Aut}(\mathrm{Fun}(G))$. The Heisenberg group collects all the commutation relations between the different group actions, and so we can say that the Heisenberg group can be presented as generated by the group actions of G, $\widehat{G}$, and $\mathrm{U}(1)$ modulo their relations in $\mathrm{Aut}(\mathrm{Fun}(G))$. For example, we recover the commutation relation $[g\cdot, \hat{g}\cdot] = \overline{\chi(g,\hat{g})}$ as the commutator of the corresponding elements in Heis:

$$\begin{aligned}[(g,1,1),(1,\hat{g},1)] &= (g^{-1},1,1)(1,\hat{g}^{-1},1)(g,1,1)(1,\hat{g},1)\\ &= (g^{-1},\hat{g}^{-1},\overline{\chi(g,\hat{g})})(g,\hat{g},\overline{\chi(g,\hat{g})})\\ &= (1,1,\overline{\chi(g,\hat{g})})\end{aligned}$$

One particular form of the Heisenberg group for $G = \mathbb{R}^n_x$ that may be familiar is as a matrix group. To make this identification, first see that the dual group $\widehat{G} = \mathbb{R}^n_t$ is really the dual vector space to G; that is, there is a natural isomorphism between $\widehat{G} = \mathrm{Hom}_{\mathbf{Group}}(\mathbb{R}^n_x, \mathrm{U}(1))$ and $G^* = \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n_x, \mathbb{R})$ that identifies $x \mapsto \exp(i\langle x,t\rangle)$ with $x \mapsto \langle x,t\rangle$ for $t \in \mathbb{R}_t$. Then the Heisenberg group for $G = \mathbb{R}_x$ (i.e., $n = 1$) is obtained from the central extension

$$1 \to \mathrm{U}(1) \to \mathrm{Heis} \to \mathbb{R}_x \times \mathbb{R}^*_x \to 0$$

Further identifications of

$$\mathrm{U}(1) \quad \text{with} \quad \left\{ \left(\begin{smallmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) \,\middle|\, z \in \mathbb{R}/\mathbb{Z} \right\}$$

and

$$\mathbb{R}_x \times \mathbb{R}^*_x \quad \text{with} \quad \left\{ \left(\begin{smallmatrix} 1 & x & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{smallmatrix}\right) \,\middle|\, x \in \mathbb{R}_x, t \in \mathbb{R}_t \right\}$$

imply that Heis can be identified with the matrix "group"

$$\left\{ \left(\begin{smallmatrix} 1 & x & z \\ 0 & 1 & t \\ 0 & 0 & 1 \end{smallmatrix}\right) \,\middle|\, x \in \mathbb{R}_x, t \in \mathbb{R}_t, z \in \mathbb{R}/\mathbb{Z} \right\}$$

where the multiplication law is matrix multiplication but we reduce modulo $\mathbb{Z}$ in the top right entry. Somehow the 2-cocycle appearing after multiplying two elements doesn't exactly match the sign conventions in the usual 2-cocycle we use for the Heisenberg group, but nevertheless we recover the commutation relation

$$\left[\left(\begin{smallmatrix} 1 & x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{smallmatrix}\right) \right] = \left(\begin{smallmatrix} 1 & 0 & -xt \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$$

Then in general the Heisenberg group for $G = \mathbb{R}^n_x$ can be thought of as the matrix "group"

$$\left\{ \left(\begin{smallmatrix} 1 & x & z \\ 0 & 1 & t \\ 0 & 0 & 1 \end{smallmatrix}\right) \,\middle|\, x \in \mathbb{R}_x, t \in \mathbb{R}_t, z \in \mathbb{R}/\mathbb{Z} \right\}$$

where the multiplication law is given by

$$\left(\begin{smallmatrix} 1 & x & z \\ 0 & 1 & t \\ 0 & 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & y & w \\ 0 & 1 & s \\ 0 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & x+y & z+w+\langle x,s\rangle \\ 0 & 1 & t+s \\ 0 & 0 & 1 \end{smallmatrix}\right)$$

where the top right entry is reduced modulo $\mathbb{Z}$. In any case, this point of view is very coordinate-ful and thus might not be a good way to think about the Heisenberg group in general.

Representation theory of the Heisenberg group

The representation theory of the Heisenberg group can be used to obtain slick proofs of results in Fourier analysis.

One important result in this vein is that $\mathrm{Fun}(G)$ is an irrep of Heis (it will not be exactly $\mathrm{Fun}(G)$ but some analysis is needed to get the function space correct; we need to have an inner product, for example, so we should think of $L^2(G)$ instead). This is to say that $\mathrm{Fun}(G)$ is a simple $\mathbb{C}[\mathrm{Heis}]$-module, and again what is meant by the group algebra is not clear, but it should include things like delta distributions and other odd elements that might not be nice functions. Contained inside this group algebra are the group algebras $(\mathbb{C}[G], *)$ and $(\mathbb{C}[\widehat{G}], *)$. The action of $\mathbb{C}[G]$ on $\mathrm{Fun}(G)$ is given by convolution; that is,

$$F \cdot f = F * f$$

for $F \in \mathbb{C}[G]$. Indeed, thinking of $g \in G$ as the element $\delta_g \in \mathbb{C}[G]$, we recover the usual action $g \cdot f = \delta_g * f = f(g^{-1}-)$. The action of $\mathbb{C}[\widehat{G}]$ on $\mathrm{Fun}(G)$ is given by first identifying $(\mathbb{C}[\widehat{G}], *)$ with $(\mathbb{C}[G], \cdot)$ by taking the inverse Fourier transform and then pointwise multiplying. Specifically this is

$$H \cdot f = (H * \hat{f})^{\vee} = H^{\vee} \cdot f$$

for $H \in \mathbb{C}[\widehat{G}]$. Indeed, thinking of $\hat{g} \in \widehat{G}$ as the element $\delta_{\hat{g}}$, we recover the usual action $\hat{g} \cdot f = (\delta_{\hat{g}} * \hat{f})^{\vee} = \chi(-, \hat{g}) f$.

To show that $\mathrm{Fun}(G)$ is an irrep of Heis we show that any nonzero function f can be sent to δ_{1_G} by acting on f by elements of the group algebra of Heis (and from the delta function we can reconstruct any other function in $\mathrm{Fun}(G)$). Indeed, given any $f \in \mathrm{Fun}(G)$ that is not the zero function, we can act on f by elements of $\mathbb{C}[\mathrm{Heis}]$ which take f to δ_{1_G}. Since $f(g) \neq 0$ for some $g \in G$, translate and scale f so that $f(1_G) = 1$. We would like to take the pointwise product $\delta_{1_G} \cdot f$ to obtain δ_{1_G}, but we obtain this product by taking the inverse Fourier transform of convolution against the character $\chi(1_G, -) \in \mathbb{C}[\widehat{G}]$. That is,

$$\chi(1_G, -) \cdot f = (\chi(1_G, -) * \hat{f})^{\vee} = \delta_{1_G} \cdot f = \delta_{1_G}$$

In general we need some analysis to approximate these delta distributions with nice enough functions to obtain the result, but this is a good enough sketch of a proof.

The action of the Heisenberg group on $L^2(\widehat{G})$ is given by swapping the roles of G and $\widehat{G}$ up to a sign since we want the action to commute with taking the Fourier transform. The element $(g, \hat{g}, z) \in \mathrm{Heis}$ acts on $f \in \mathrm{Fun}(G)$ by

$$(g, \hat{g}, z) \cdot f = z\chi(-, \hat{g}) f(g^{-1}-)$$

and on $F \in \mathrm{Fun}(\widehat{G})$ by

$$(g, \hat{g}, z) \cdot F = z\chi(g, \hat{g})\chi(g, -)F(\hat{g}-)$$

(this is just interchanging the roles of G and $\widehat{G}$ in some sense). Then the Fourier transform (and hence also the inverse Fourier transform) intertwines with this action:

$$\widehat{(g, \hat{g}, z) \cdot f} = z\widehat{\chi(g, -)\delta_g * f} = z\delta_{\hat{g}^{-1}} * (\chi(g, -)\hat{f}) = z\chi(g, \hat{g})\chi(g, -)\hat{f}(\hat{g}-) = (g, \hat{g}, z) \cdot \hat{f}$$

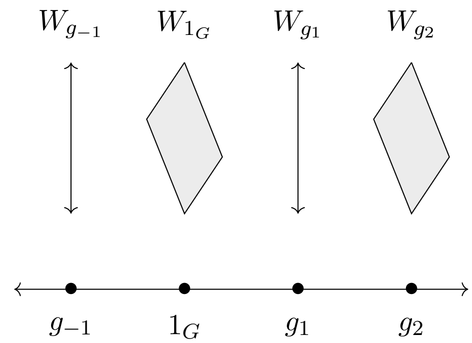
Since this action is not really any different than the original action we defined, $\mathrm{Fun}(\widehat{G})$ is also an irrep of Heis (and we can repeat this as well for $\mathrm{Fun}(\widehat{\widehat{G}})$).

Let $G = \mathbb{R}_x$. That the Fourier transform is a unitary isomorphism of $L^2(G)$ with $L^2(\widehat{G})$ can also be proved knowing some kind of unitary Schur's lemma, which comes from analysis. Identify G with $\widehat{G}$; further identify $L^2(G)$ with $\overline{L^2(G)}$ and $L^2(G)^*$ simultaneously. The argument in analysis is that the Fourier transform is densely defined and can be closed (we can take limits of Fourier transforms), so its adjoint is also densely defined, which implies the Fourier transform must be unitary up to a constant multiple. Morally what is happening is that intertwining operators of unitary representations are typically unitary operators (they would "have to be terrible not to be unitary" according to Howe). This is usually known as Plancherel's theorem. (See p.452 in Real and Functional Analysis by Lang.) In other words, the Fourier transform takes an invariant Hermitian form $\langle -, - \rangle$ on $L^2(G)$; that is, an isomorphism of Heis-representations $\overline{L^2(G)} \to L^2(G)^*$ to another invariant Hermitian form $\langle \widehat{-}, \widehat{-} \rangle$ on $L^2(G)$; that is, to another isomorphism $\overline{L^2(G)} \to L^2(G)^*$. But by Schur's lemma these two forms should only differ by a constant multiple and to find out which multiple it suffices to calculate the new Hermitian form on a convenient element to see that the Fourier transform is unitary.

To prove that $\hat{\hat{f}}(g) = f(g^{-1})$, observe that applying the Fourier transform twice intertwines with correctly chosen actions of the Heisenberg groups. The reflection map $\mathrm{Fun}(G) \xrightarrow{R} \mathrm{Fun}(G)$ given by $Rf = f((-)^{-1})$ also intertwines with the action of the Heisenberg group, so applying the Fourier transform twice is a scalar multiple of R. To find out what the multiple is, it suffices to calculate the twice Fourier transform of a single element; usually we pick a convenient element. In the case when $G = \mathbb{R}$, the Gaussian $f(x) = \exp\big(-x^2/2\big)$ is self-dual, meaning $\hat{f}(t) = \exp\big(-t^2/2\big)$ from which it follows that taking the Fourier transform twice really does agree with the reflection map. But this is not too bad to do since f satisfies the differential equation $\frac{\mathrm{d}}{\mathrm{d}x}u + xu = 0$. The Fourier transform of this differential equation is $-it\hat{u} - i\frac{\mathrm{d}}{\mathrm{d}t}\hat{u} = 0$; that is, $\frac{\mathrm{d}}{\mathrm{d}t}\hat{u} + t\hat{u} = 0$, from which it follows that the Fourier transform of the aforementioned Gaussian is itself.

One of the most important results about the Heisenberg group is the Stone-von Neumann theorem, which roughly says that $\mathrm{Fun}(G)$ is the only genuine irrep of the Heisenberg group up to isomorphism, where by a genuine representation V we mean that the subgroup $\mathrm{U}(1) \subset \mathrm{Heis}$ acts on V by the usual scalar multiplication of complex numbers on V.

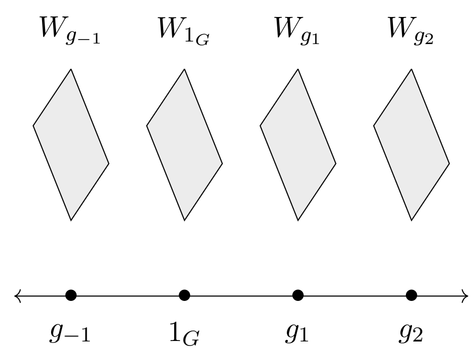
A stronger result is that the categories $\mathbf{Rep}_{\text{genuine}}(\mathrm{Heis})$ and $\mathbf{Vect}_{\mathbb{C}}$ are equivalent, which implies the Stone-von Neumann theorem since in this setting $\mathrm{Fun}(G)$ corresponds to the one-dimensional vector space $\mathbb{C}$. Let W be a genuine representation of Heis. Restricting to the action of $G, \widehat{G}$ in Heis, view W as a representation of $\widehat{G}$, which from before it follows that W can be thought of as global sections of a sheaf over $\widehat{\widehat{G}} \cong G$. The following picture is slightly inaccurate but we will fix it shortly: (The inaccuracy is that the vector spaces W_g have different dimensions; the picture is an example where the drawn vector spaces have dimension 1 or 2.)



On W_g, the group $\widehat{G}$ acts by multiplication by $\chi(g, \hat{g})$. Think of W_g as the $\chi(g, -)$-eigenspace for the $\widehat{G}$-action on W. On the other hand, G acts on $w \in W_g$ by translation, specifically by translating the eigenspace w belongs to. That is, if $h \in G$ and $w \in W_g$, then by the commutation relation of G and $\widehat{G}$ in Heis we have

$$\hat{g}(hw) = (\hat{g}h)w = \chi(h, \hat{g})(h\hat{g})w = \chi(hg, \hat{g})(hw)$$

So $h \in G$ sends $w \in W_g$ to $hw \in W_{hg}$. Since multiplication by group elements in G is invertible, the eigenspaces W_g are all isomorphic to each other (i.e. for any $h, g \in G$, $W_g \cong W_{hg}$ as $\widehat{G}$-representations), so the picture from earlier is more accurately depicted with each stalk the same dimension.



The equivalence $\mathbf{Rep}_{\text{genuine}}(\text{Heis}) \to \mathbf{Vect}_{\mathbb{C}}$ is given by taking the stalk of the sheaf we get by smearing a representation W at 1_G. The functor going the other way is given by taking a complex vector space V, letting $\widehat{G}$ act on it by multiplication by multiplication by $\chi(1_G, -)$; that is, thinking of V as the stalk of a sheaf at 1_G. Then "translate" $V_{1_G} = V$ by G to form copies V_g for which $\widehat{G}$ acts by multiplication by $\chi(g, -)$. The global section of the sheaf we obtain by translating V this way is a genuine representation of Heis if throughout we let $\mathrm{U}(1)$ act by usual scalar multiplication. The only way to obtain an irrep of Heis through this functor is by starting with $V = \mathbb{C}$, which when translated over G yields a sheaf whose global sections is exactly $\mathrm{Fun}(G)$.

So the role of $\widehat{G}$ in this picture is to spectrally decompose a representation W and G translates the eigenspaces around. Of course, there is a symmetric picture if we swap the roles of G and $\widehat{G}$ around.

In physics, we look at the infinitesimal actions of translation from $G = \mathbb{R}_x$ and multiplication from $\widehat{G} = \mathbb{R}_t \cong \mathbb{R}_x$. From some calculations from earlier, we saw that these correspond to $\frac{\mathrm{d}}{\mathrm{d}x}$ and x; the observables of position and momentum (in physics some additional constants are present). These two operators do not commute, and hence cannot be simultaneously diagonalized, which is the content of the Heisenberg uncertainty principle. The Stone-von Neumann theorem says that this is pretty much the only case up to scale.

A short preview of the Weyl algebra

Let $G = \mathbb{R}_x$. The Weyl algebra is a non-commuting $\mathbb{C}$-algebra of polynomial differential operators:

$$D = \frac{\mathbb{C}\langle x, \partial_x \rangle}{(\partial_x x - x\partial_x - 1)}$$

The Weyl algebra naturally acts on $\mathrm{Fun}(G)$ or rather $L^2(G)$ by unbounded operators, and Hermann Weyl used this algebra and related algebras to study the Heisenberg uncertainty principle in quantum mechanics. It is related to the Heisenberg group from earlier since differentiation and multiplication by x are the infinitesimal actions of $G, \widehat{G}$ on $L^2(G)$, but there are very interesting differences in the theories surrounding representations of the Weyl algebra and of the Heisenberg group.

The Weyl algebra appears in the theory of D-modules, which from one point of view is a way to study the theory of systems of linear partial differential equations using algebra and algebraic geometry. It is (one of) David's favorite algebras.

Lecture 8 September 18

The Heisenberg group and the Weyl algebra

The Heisenberg group is not a semidirect product of U(1) and $G \times \widehat{G}$, but by breaking the symmetry of the pair $G, \widehat{G}$, it is a semidirect product of $\widehat{G} \times \mathrm{U}(1)$ and G. That is,

$$\mathrm{Heis} \cong G \ltimes_\varphi \widehat{G} \times \mathrm{U}(1)$$

where $\varphi\colon G \to \mathrm{Aut}(\widehat{G} \times \mathrm{U}(1))$ is given by $\phi(g)(\hat{h}, w) = (\hat{h}, \overline{\chi(g, \hat{h})}w)$. Hence the group multiplication is given by

$$(g, \hat{g}, z)(h, \hat{h}, w) = (gh, \hat{g}\hat{h}, \overline{\chi(g, \hat{h})}zw)$$

and agrees with the usual one on Heis.

The observations we made last time about viewing genuine representations of Heis as a sheaf over G in which $\widehat{G}, \mathrm{U}(1)$ act in a manner that preserve stalks and G acts by moving vectors between stalks can be recast using the above isomorphism. Genuine irreps of $\widehat{G} \times \mathrm{U}(1)$ of a fixed dimension are in correspondence with (or live over) the elements of G, and G acts on the set of irreps by changing the action of $\widehat{G}$ on them; that is, by exchanging irreps (last time, this was to exchange W_g with W_{hg}).

Last time, we saw that the Weyl algebra D on $\mathbb{R}_x$ was the noncommutative $\mathbb{C}$-algebra generated by x and ∂_x, subject to the relation $\partial_x x - x\partial_x = 1$. It acts on $L^2(\mathbb{R})$, but a better space for the Weyl algebra to act on is the Schwartz space $\mathcal{S}(\mathbb{R})$. The Schwartz space contains the C^∞ functions on $\mathbb{R}$ whose derivatives of all orders decay faster than the reciprocal of any polynomial as x tends to $\pm\infty$. A precise definition of the Schwartz space and its topology is rather cumbersome, so it suffices to think of the functions living in the Schwartz space as "rapidly decaying smooth functions". For example, the Gaussian $\exp(-x^2/2)$ belongs to $\mathcal{S}(\mathbb{R})$.

By design, the Schwartz space is a module for the Weyl algebra, since the functions in the Schwartz space and their derivatives rapidly decrease, multiplication by x and differentiation really define an action on $\mathcal{S}(\mathbb{R})$. By taking the continuous dual of the Schwartz space, we obtain the tempered distributions, denoted $\mathcal{S}'(\mathbb{R})$. The tempered distributions $\mathcal{S}'(\mathbb{R})$ is another candidate for the group algebra for $G = \mathbb{R}_x$.

From analysis, the Fourier transform defines operators on $L^2(\mathbb{R})$ and on $\mathcal{S}(\mathbb{R})$. By taking adjoints, the Fourier transform also defines an operator on $\mathcal{S}'(\mathbb{R})$. So the space of tempered distributions is a very nice place to do harmonic analysis on, and we would like to do something similar in the future for non-Abelian groups.

There are no nonzero finite-dimensional D-modules. Suppose V is a finite-dimensional D-module. Then

$$\dim V = \mathrm{Tr}_V(1) = \mathrm{Tr}_V(\partial_x x - x\partial_x) = \mathrm{Tr}_V(\partial_x x) - \mathrm{Tr}_V(x\partial_x) = \mathrm{Tr}_V(\partial_x x) - \mathrm{Tr}(\partial_x x) = 0$$

This result may be thought of as an algebraic version of the Heisenberg uncertainty principle, since not only can we not simultaneously diagonalize ∂_x and x, but we cannot have these operators act on finite dimensional vector spaces.

The exponentials $\exp(iy-)$ of $i\partial_x$ and of x give the translation τ_y and multiplication by the character $\exp(iyx)$, and these generate the Heisenberg group. By starting with a D-module, we can obtain a module over Heis by differentiating the action of $G, \widehat{G}$ to obtain differentiation and multiplication by x. To be careful, we need analysis because D-modules are not finite-dimensional.

Quick introduction to D-modules

The data of a finitely presented D-module is the same as a system of linear differential equations with polynomial coefficients. We will come back to this claim after looking at small examples.

Let $M = D/D(\partial_x - \lambda)$ be a left D-module (left, right are important since D is not commutative). Then for some other left D-module F, which is typically a nice function space, $f \in \mathrm{Hom}_D(M, F)$ can be thought of as a solution to the differential equation $(\partial_x - \lambda)u = 0$ living in F. This is because any $f \in \mathrm{Hom}_D(M, F)$ is zero on $D(\partial_x - \lambda)$ and any $f \in \mathrm{Hom}_D(D, F)$ which is zero on $D(\partial_x - \lambda)$ descends to a well-defined homomorphism $f \in \mathrm{Hom}_D(M, F)$. That is,

$$\mathrm{Hom}_D(M, F) \cong \{f \in \mathrm{Hom}_D(D, F) \mid f(\partial_x - \lambda) = 0\}$$

The map $f \mapsto f(1_M)$ defines an isomorphism of $\mathrm{Hom}_D(D, F)$ with F, so in particular

$$\{f \in \mathrm{Hom}_D(D, F) \mid f(\partial_x - \lambda) = (\partial_x - \lambda)f(1_M) = 0\} \cong \{f \in F \mid (\partial_x - \lambda)f = 0\}$$

So if F is a D-module that contains $\exp(\lambda x)$, then $\mathrm{Hom}_D(M, F)$ is $\mathbb{C}\exp(\lambda x)$, otherwise $\mathrm{Hom}_D(M, F)$ is zero. One example of a D-module F that has exponentials is the continuous dual of $C_c^\infty(\mathbb{R})$, the usual space of distributions on $\mathbb{R}$ (analysis shows that $\exp(\lambda x)$ does not define a tempered distribution on $\mathbb{R}$). As a vector space (or $\mathbb{C}[x]$-module), the module M is isomorphic to the graded vector space $\mathbb{C}[x]1_M$. The calculation $(\partial_x - \lambda)x = 1 - x\partial_x - x\lambda = 1$ shows that left multiplication by $(\partial_x - \lambda)$ gives maps between graded components of $\mathbb{C}[x]1_M$:

$$\begin{tikzcd} 0 \\ \mathbb{C}1_M \arrow[u, bend left, "(\partial_x-\lambda)"] \\ \mathbb{C}x \arrow[u, bend left, "(\partial_x-\lambda)"] \\ \vdots \arrow[u, bend left, "(\partial_x-\lambda)"] \end{tikzcd}$$

This graded point of view is interesting because now we can see that if $f \in F$ is a solution to $(\partial_x - \lambda)u = 0$; that is, f is a homomorphism of D-modules $D/D(\partial_x - \lambda) \to F$, then $x^n f(x)$ is a solution to $(\partial_x - \lambda)^{n+1}u = 0$. This is because $(\partial_x - \lambda)^{n+1}x^n = 0$ in M. So in particular, we recover the expected solutions $x^n \exp(\lambda x)$.

Consider the left D-module $M = D/D(x - \lambda)$. Its solutions in $\mathcal{S}'(\mathbb{R})$ are $\mathbb{C}\delta_\lambda$. As a vector space (or $\mathbb{C}[\partial_x]$-module), the module M is isomorphic to the graded vector space $\mathbb{C}[\partial_x]1_M$. We obtain a similar picture as above, with x

and ∂_x interchanged. Similarly, solutions to the equation $(x-\lambda)^n u = 0$ are derivatives of δ_λ. We actually already observed these facts when discussing the indecomposable representations of $\mathbb{R}$, from the point of view of Jordan blocks and Fourier duality, so this is just another perspective.

A finitely presented left D-module M is a module with the free resolution

$$0 \to D^n \xrightarrow{P} D^m \to M \to 0$$

where m denotes the number of generators of M and the map P, which can be thought of as right multiplication by an $n \times m$ matrix P (with entries in D), encodes the n relations that the m generators of M are subject to. In other words, M is isomorphic to the left D-module D^m/D^nP. Like before, if F is a left D-module, then $f \in \mathrm{Hom}_D(M,F)$ can be identified with solutions $f = (f_1, \ldots, f_m)$ to the system of differential equations

$$\sum_{i=1}^n P_{ij} f_j = 0 \quad \text{for } j = 1, \ldots, m$$

This is because $f = (f_1, \ldots, f_m) \in \mathrm{Hom}_D(M,F)$ is really a homomorphism $f = (f_1, \ldots, f_m)$ from D^m to F vanishing on $1_{D^n}P = (\sum_{i=1}^n P_{i1}, \ldots, \sum_{i=1}^n P_{im})$, and we identify $\mathrm{Hom}_D(D,F)$ with F in the same way as before.

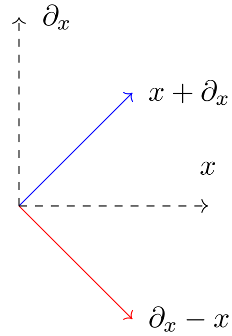
If $G = \mathbb{R}^n$ (with coordinates $x_1, \ldots, x_n$), then the Weyl algebra D takes on the form

$$D = \frac{\mathbb{C}\langle x_1, \ldots, x_n, \partial_{x_1}, \ldots, \partial_{x_n}\rangle}{([x_i, x_j], [\partial_{x_i}, \partial_{x_j}], [\partial_{x_i}, x_j] - \delta_{ij} \mid 1 \le i, j \le n)}$$

It follows that the data of a finitely presented D-module is the same as a system of linear partial differential equations with polynomial coefficients, which we can see by essentially repeating the above argument.

The Fourier transform takes $\frac{\mathrm{d}}{\mathrm{d}x}$ to $-it$ and x to $-i\frac{\mathrm{d}}{\mathrm{d}t}$. It follows that $1 = \partial_x x - x\partial_x$ is sent to $1 = (-it)(-i\partial_t) - (-i\partial_t)(-it) = \partial_t t - t\partial_t$, so the Fourier transform negates the commutation relation between differentiation and multiplication by the coordinate. Since $\mathcal{S}'(\mathbb{R})$ is a D-module, the Fourier transform of $\mathcal{S}'(\mathbb{R})$, which is itself $\mathcal{S}'(\mathbb{R})$, is a $\widehat{D} = \mathbb{C}\langle t, \partial_t\rangle/([\partial_t, t] + 1)$-module (a module for some kind of opposite version of D). So if f satisfies a linear differential equation with polynomial coefficients $Pu = 0$ for a left D-module M, then $\hat{f}$ satisfies the Fourier transformed partial differential equation $\hat{P}\hat{u} = 0$, so for the left D-module $\widehat{M}$ which is M with ∂_x, x replaced by $-it, -i\partial_t$.

Indeed, the Gaussian $g = \exp\bigl(-x^2/2\bigr)$ satisfies the differential equation $(\partial_x + x)u = 0$. It defines a tempered distribution as well; that is, $g \in \mathcal{S}'(\mathbb{R})$. The Fourier transform takes this differential equation to $(-it - i\partial_t)\hat{u} = 0$, so $(\partial_t + t)\hat{u} = 0$. So the Fourier transform of the Gaussian is itself (we did this calculation before). Consider the element $(\partial_x - x)$ in M. Suggestively, it is "orthogonal" to the element $(x + \partial_x)$:



The corresponding D-module for g is $M = D/D(\partial_x + x)$, which is isomorphic as a vector space to $\mathbb{C}[x - \partial_x]1$ (the $(x - \partial_x)^i$ form a basis for M). The calculation $(\partial_x + x)(x - \partial_x) = 1 + x^2 - \partial_x^2 = 2 + (x - \partial_x)(\partial_x + x) = 2$ shows that left multiplication by $(\partial_x + x)$ gives maps between graded components of $\mathbb{C}[x - \partial_x]1$:

$$\begin{tikzcd} 0 \\ \mathbb{C}1 \arrow[u, bend left, "(\partial_x+x)"] \\ \mathbb{C}(x - \partial_x) \arrow[u, bend left, "(\partial_x+x)"] \\ \vdots \arrow[u, bend left, "(\partial_x+x)"] \end{tikzcd}$$

The homomorphism $S_g \colon M \to \mathcal{S}'(\mathbb{R})$ corresponding to the solution g of the differential equation $(\partial_x + x)u = 0$ is the one where $1 \in M$ is sent to $g \in \mathcal{S}'(\mathbb{R})$. Then $(x - \partial_x)$ is sent to $(x - \partial_x)g = 2xg$, and $(x - \partial_x)^2$ is sent to $(x - \partial_x)^2 g = (4x^2 - 2)g$. By continuing this process we recover the physicist's Hermite polynomials $H_n(x)$: the image of $(x - \partial_x)^n$ is $H_n(x)g$, where the first few Hermite polynomials are

$$\begin{aligned} H_0(x) &= 1 & H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 & H_3(x) &= 8x^3 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 & H_5(x) &= 32x^5 - 160x^3 + 120x \end{aligned}$$

The Hermite polynomials, like the Chebyshev polynomials of the first kind we found earlier, are special functions. They satisfy nice orthogonality relations and recurrence relations which can of course be obtained through the lens of representation theory. The Hermite polynomials form a basis of $L^2(\mathbb{R})$ with weighted inner product $\langle -, - \rangle_g = \langle -g, -g \rangle_{L^2(\mathbb{R})}$ because they form a maximal orthogonal family. To see this, observe that

$$(\partial_x + x)(x - \partial_x) = 2 + (x - \partial_x)(\partial_x + x)$$

and that the adjoint differential operator of $(x - \partial_x)$ is $(\partial_x + x)$; that is,

$$\langle (x - \partial_x)f, h \rangle_{L^2(\mathbb{R})} = \langle f, (\partial_x + x)h \rangle_{L^2(\mathbb{R})}$$

Repeated application of these identities will show that

$$\langle H_j(x), H_i(x)\rangle_g = \langle H_j(x)g, H_i(x)g\rangle_{L^2(\mathbb{R})} = \langle (x-\partial_x)^j g, (x-\partial_x)^i g\rangle_{L^2(\mathbb{R})}$$

is nonzero only if $j = i$.

Introduction to SU(2)

The main groups we would like to get to are SU(2), $\mathrm{SO}_3(\mathbb{R})$, and of course, $\mathrm{SL}_2(\mathbb{R})$. By studying the representation theory of the first two groups (which are subgroups of $\mathrm{SL}_2(\mathbb{R})$), we will gain some insight for how to study the representation theory of $\mathrm{SL}_2(\mathbb{R})$.

By identifying complex numbers $z = x + iy$ with real matrices $\left(\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right)$, we can identify $\mathrm{U}(1) = \{z \in \mathbb{C} \mid |z| = 1\}$ with the real matrix group $\{\left(\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right)\} \cap \mathrm{SL}_2(\mathbb{R})$. There is a natural complexification of this idea (it's unclear whether this was a pun or not).

One way to describe the quaternions is as a complex matrix group; that is,

$$\mathbb{H} = \{\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)\}$$

viewed as a subgroup of the 2×2 complex matrices. The quaternions form an associative, non-commutative $\mathbb{C}$-algebra which may also be described as an $\mathbb{R}$-algebra spanned by $\{1, i, j, k\}$ subject to the relations $i^2 = j^2 = k^2 = -1$, $ij = k$, $jk = i$, $ki = j$. The description of $\mathbb{H}$ as an $\mathbb{R}$-algebra is convenient since there is a nice way to write the conjugation involution and norm on $\mathbb{H}$: For $a + bi + cj + dk \in \mathbb{H}$, $\overline{a + bi + cj + dk} = a - bi - cj - dk$ and $\|a + bi + cj + dk\|^2 = (a + bi + cj + dk)(\overline{a + bi + cj + dk}) = a^2 + b^2 + c^2 + d^2$. We can also view the quaternions as the complex matrix algebra generated by the matrices $1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $i = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, $j = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $k = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$.

In the same way that we think of U(1) as the determinant 1 matrices in $\mathrm{SL}_2(\mathbb{R})$, we can define SU(2) as the unit quaternions; that is,

$$\mathrm{SU}(2) = \mathbb{H} \cap \mathrm{SL}_2(\mathbb{C})$$

The determinant 1 matrices $\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)$ are those with $|z|^2 + |w|^2 = 1$; if $z = z_1 + iz_2$ and $w = w_1 + iw_2$, then $z_1^2 + z_2^2 + w_1^2 + w_2^2 = 1$. So the elements of SU(2), the unit quaternions, can be thought of the 3-sphere S^3.

Related is the group of 2×2 unitary matrices U(2), whose elements satisfy $U\overline{U}^\top = \overline{U}^\top U = 1$ (so $|\det U| = 1$). The group SU(2) is a subgroup of U(2), and we can also write $\mathrm{SU}(2) = \mathrm{U}(2) \cap \mathrm{SL}_2(\mathbb{C})$.

That we can identify the unit quaternions with S^3 shows that we can endow S^3 with a group structure. Similarly, the unit magnitude elements of $\mathbb{C}$ may be identified with S^1 to give S^1 a group structure, and the same can be said about S^0 and the unit magnitude elements $\{-1, 1\}$ of $\mathbb{R}$.

Some easy complex representations of SU(2) are $\mathbb{H} \cong \mathbb{C}^2$, one where SU(2) acts by left multiplication and another where SU(2) acts by conjugation. The first representation is irreducible, since any element of a proper subspace of

$\mathbb{H}$ can be translated out of that subspace by a suitably chosen element of SU(2). The conjugation representation viewed as a real representation is reducible, since by viewing $\mathbb{H}$ as $\mathbb{R}^4$, the subspace $\mathbb{R}1$ is preserved under conjugation by unit quaternions. It is also the case that the complement of $\mathbb{R}1$, the pure quaternions (which we identify as $\mathbb{R}^3$ so that $\mathbb{H} = \mathbb{R}1 \oplus \mathbb{R}^3$ as a real vector space), is also preserved under conjugation by unit quaternions. This is because for a unit quaternion q and quaternions g, h, the real part of h is $(h + \overline{h})/2$ and the pure quaternion part of h is $(h - \overline{h})/2$, conjugation of a product is order-reversing: $\overline{gh} = \overline{h}\overline{g}$, and conjugation by q commutes with conjugation $q\overline{h}q^{-1} = q\overline{h}\overline{q} = \overline{q h \overline{q}} = \overline{qhq^{-1}}$.

That the pure quaternions are a 3-dimensional real representation of SU(2) under conjugation is to say there is a particular homomorphism $\mathrm{SU}(2) \to \mathrm{GL}(\mathbb{R}^3)$ which manifests the conjugation action. Conjugation by (nonzero) quaternions preserves the norm of any quaternion (it is easier to see this by returning to the matrix form of quaternions and observing that conjugation by a quaternion preserves the determinant, since the determinant of a matrix is invariant under conjugation). A calculation (not shown here, but see Wikipedia) shows that the orientation of a pure quaternion viewed as an element of $\mathbb{R}^3$ is also preserved under conjugation. It follows that the image of the homomorphism SU(2) to $\mathrm{GL}(\mathbb{R}^3)$ is a subgroup of $\mathrm{SO}_3(\mathbb{R})$. This amounts to viewing conjugation of pure quaternions by a unit quaternion as rotations of vectors in $\mathbb{R}^3$.

The map from SU(2) to $\mathrm{SO}_3(\mathbb{R})$ is actually a surjective two-to-one map with kernel $\{\pm 1\} \cong \mathbb{Z}/2\mathbb{Z}$. So we get an exact sequence

$$1 \to \{\pm 1\} \to \mathrm{SU}(2) \to \mathrm{SO}_3(\mathbb{R}) \to 1$$

where $\mathrm{SO}_3(\mathbb{R}) = \mathrm{SU}(2)/\{\pm 1\}$. By identifying SU(2) with S^3, quotienting out by $\{\pm 1\}$ amounts to identifying antipodal points of S^3, so as a topological space, $\mathrm{SO}_3(\mathbb{R})$ is identified with $\mathbb{RP}^3$. We say that SU(2) is a double cover of $\mathrm{SO}_3(\mathbb{R})$.

David brought an interesting prop made of a belt and a piece of cardboard that represents the double cover of $\mathrm{SO}_3(\mathbb{R})$ by SU(2). The belt is supposed to keep track of the twisting by $\{\pm 1\}$ and the cardboard is supposed to keep track of the position of an element of $\mathrm{SO}_3(\mathbb{R})$. I would like to make a video about this that I can link here.

A fun fact: Let u be a pure quaternion of unit length and let $g_\theta = \cos(\theta/2) + \sin(\theta/2)u$. Then g_θ is a unit quaternion and conjugation of a pure quaternion v by g_θ (or $-g_\theta$) can be viewed in $\mathbb{R}^3$ (viewing u, v as elements of $\mathbb{R}^3$) as rotating v by θ radians about the axis u (i.e. counterclockwise about u if we move our point of view so that u points towards us).

By letting θ vary, we can identify $\{g_\theta\} \subset \mathrm{SU}(2) = S^3$ with a copy of $\mathrm{U}(1) = S^1$. Then we can draw the commutative cube

$$\begin{tikzcd}[row sep=small, column sep=small]
& \mathbb{C}^\times \arrow[rr] \arrow[dd] & & \mathrm{SL}_2(\mathbb{C}) \arrow[dd] \\
\mathrm{U}(1) \cong S^1 \arrow[ur] \arrow[rr, crossing over] \arrow[dd] & & \mathrm{SU}(2) \cong S^3 \arrow[ur] & \\
& \mathbb{C}^\times/\{\pm 1\} \arrow[rr] & & \mathrm{SL}_2(\mathbb{C})/\{\pm 1\} \cong \mathrm{PSL}_2(\mathbb{C}) \\
\mathrm{U}(1)/\{\pm 1\} \arrow[ur] \arrow[rr] & & \mathrm{SO}_3(\mathbb{R}) \cong \mathbb{RP}^3 \arrow[ur] \arrow[from=uu, crossing over] &
\end{tikzcd}$$

We will try to study the representation theory of $\mathrm{SU}(2)$ and $\mathrm{SO}_3(\mathbb{R})$ by first using the representation theory of $\mathrm{U}(1)$, which we have already studied in detail.

Lecture 9 September 23

Prologue to representation theory of SU(2)

There are several ways to study representations of SU(2), but the approach we will take is to start by first using the representation theory of U(1) to extract interesting information about representations of SU(2) as opposed to constructing them explicitly via a robust structure theory or some other approach.

Let V be a finite dimensional complex (continuous) representation of SU(2). Since SU(2) is not Abelian, we do not immediately get to use the nice theory of LCA groups we discussed earlier. So instead we look at the action of a copy of U(1) living inside SU(2) on V; that is, we think of V first as a U(1)-representation. There are lots of different copies of U(1) inside of SU(2), but for now think of the copy given by $\{\cos(\theta)+\sin(\theta)i\} \subset \mathrm{SU}(2)$ (the unit quaternions where the coefficients of j, k are zero). A copy of U(1) in SU(2) is called a (maximal) torus, usually denoted T (a torus in a compact Lie group G is a compact, connected, Abelian Lie subgroup of G, and a maximal torus is a torus that is maximal under inclusion of tori).

From the earlier U(1)-theory, V is unitarizable under the action of T and hence decomposes into isotypic components $V = \bigoplus_{n\in\mathbb{Z}} V_n$ where each V_n is $\dim V_n$ many copies of the n-th irrep $\mathbb{C}_{\chi_n}$ of T where $z \cdot w = \chi_n(z)w$ where χ_n is a character of T; that is,

$$V_n = \mathbb{C}^{\dim V_n} \otimes_{\mathbb{C}} \mathbb{C}_{\chi_n}$$

In other words, V_n keeps track of the possibly repeated one-dimensional irreps. All but finitely many $\dim V_n$ are zero since V is finite dimensional. This decomposition of V is called the weight space decomposition with respect to the maximal torus T inside SU(2), and the V_n are called the weight spaces.

One function we can get from this decomposition is a function $\mathbb{Z} \to \mathbb{Z}$ given by $n \mapsto \dim V_n$ (which is of course finitely supported since V is finite dimensional). From this function form the Laurent polynomial $\chi_V(z)$ given by

$$z \mapsto \chi_V(z) = \sum_{n\in\mathbb{Z}} (\dim V_n) z^n \in \mathbb{C}[z, z^{-1}]$$

If we use the symbol z for an element $\exp(i\theta) \in \mathrm{U}(1) \cong T$ (identifying z with $\left(\begin{smallmatrix} z & 0 \\ 0 & \bar{z} \end{smallmatrix}\right)$), then $z \cdot v$ for $v \in V_n$ is $z^n v$. The $\dim V_n \times \dim V_n$ matrix for the action of z on $V_n = \mathbb{C}^{\dim V_n} \otimes_{\mathbb{C}} C_{\chi_n}$ is given by

$$z \cdot v = \begin{pmatrix} z^n & & \\ & \ddots & \\ & & z^n \end{pmatrix} v$$

Thus if $z \in \mathrm{U}(1)$, then $\mathrm{Tr}_{V_n}(z\cdot) = (\dim V_n) z^n$; similarly, $\chi_V(z) = \sum_{n\in\mathbb{Z}} (\dim V_n) z^n$ is the trace $\mathrm{Tr}_V(z\cdot)$. So think of the Laurent polynomial $\chi_V(z)$ as a function $\mathrm{U}(1) \to \mathbb{C}$, which we call the character of the U(1)-representation V.

But we don't need to restrict χ_V to T; since V is finite-dimensional, we could instead define more generally define the character χ_V on G by $\chi_V(g) = \mathrm{Tr}_V(g\cdot)$; that is, by taking the trace of the group action on V. Another observation is that the character χ_V is invariant under conjugation since the trace is: $\chi_V(hgh^{-1}) = \mathrm{Tr}_V(hgh^{-1}\cdot) = \mathrm{Tr}_V(g\cdot) = \chi_V(g)$. Functions that are invariant under conjugation are called class functions. The approach we are heading towards to study the representation theory of SU(2) is called character theory.

The Weyl group

Let V be an irrep of a group G. By Schur's lemma $Z(G) \subset G$ acts on V by scalar multiples of id_V since acting on V by an element of $Z(G)$ defines an intertwining operator $V \to V$ (because elements of $Z(G)$ commute with every element of G).

The center of SU(2) is $\{\pm 1\}$: any matrix $\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)$ in the center of SU(2) must satisfy $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)$, so we observe first that $\operatorname{Im} z = 0$ and $\operatorname{Re} w = 0$, then from $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)\left(\begin{smallmatrix} z & w \\ w & z \end{smallmatrix}\right)\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} z & w \\ w & z \end{smallmatrix}\right)$, deduce that $\operatorname{Im} w = 0$. Then $\det\left(\begin{smallmatrix} z & 0 \\ 0 & z \end{smallmatrix}\right) = 1$ so that $z = \pm 1$. As a side note, it follows that SU(2) is an extension of $\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}_3(\mathbb{R})$ by $Z(\mathrm{SU}(2)) = \{\pm 1\} \cong \mathbb{Z}/2\mathbb{Z}$, which is central but non-split. This is because a section $\mathrm{SO}_3(\mathbb{R}) \to \mathrm{SU}(2)$ must be continuous and hence preserve loops. Consider the element in $\mathrm{SO}_3(\mathbb{R})$ given by fixing an axis $u \in \mathbb{R}^3$ (with u of unit length) and rotating counterclockwise by an angle θ, denote this element by $R_{u,\theta}$. The element $R_{u,\theta}$ should be sent to the element $g_\theta = \cos(\theta/2) + u\sin(\theta/2)$ in SU(2) (here we are identifying the unit pure quaternions with $S^2 \subset \mathbb{R}^3$). So when $\theta = 0$, $g_\theta = 1$ and when $\theta = 2\pi$, $g_\theta = -1$. But the collection $\{R_{u,\theta} \mid 0 \le \theta \le 2\pi\}$ forms a loop in $\mathrm{SO}_3(\mathbb{R})$, but this loop is not sent to a loop in SU(2). So no section can exist.

The observations of the previous two paragraphs combine to reveal that $-1 \in \mathrm{SU}(2)$ acts on an irrep V of SU(2) by a scalar multiple of id_V. Since $-1^2 = 1$, we are limited to the cases that either -1 can act as $\mathrm{id}\, V$ or $-\mathrm{id}_V$. Sort irreps of SU(2) by whether -1 acts trivially by id_V or "genuinely" by $-\mathrm{id}_V$. The irreps in the first collection are called even irreps of SU(2), and the irreps in the second collection are called odd irreps of SU(2). The even irreps of SU(2) may be viewed as an irrep of $\mathrm{SO}_3(\mathbb{R})$, since last time we saw that $\mathrm{SO}_3(\mathbb{R}) \cong \mathrm{SU}(2)/\{\pm 1\}$ (if both ± 1 act trivially on an even irrep V of SU(2), then the action of the quotient $\mathrm{SU}(2)/\{\pm 1\}$ on V is well-defined and V will remain irreducible under the new action).

If V is an even irrep of SU(2), its decomposition as a U(1)-representation is $V = \bigoplus_{n\in\mathbb{Z}} V_n$ where $\dim V_n$ for n odd is zero. This is because -1 is contained in U(1), and the the action of -1 on $\mathbb{C}_{\chi_n}$ is given by scalar multiplication by $(-1)^n$, so in order for the action of -1 on V_n to be trivial for odd n, $\dim V_n$ must be zero.

Consider a representation V of SU(2). As before, the weight space decomposition of V with respect to T is $V = \bigoplus_{n\in\mathbb{Z}} V_n$. Let h be an element of the normalizer $N(T) \subset \mathrm{SU}(2)$ of the torus T. Observe that $\chi_n(h^{-1}(-)h)$ defines a character of the torus T, which we know is of the form χ_m for some m (where it is of course possible that $m = n$).

Then for any $t \in T$ and $v \in V_n$,

$$t \cdot (h \cdot v) = th \cdot v = hh^{-1}th \cdot v = h \cdot (\chi_n(h^{-1}th)v) = \chi_m(t)(h \cdot v)$$

This shows that the action of the normalizer of T on elements of the weight spaces V_n is to possibly move elements between weight spaces. Furthermore, $h \cdot V_n$ is isomorphic to V_n since multiplication by h is invertible. So the action of h on the weight spaces is a permutation of the weight spaces. Observe that the action of $T \subset N(T)$ does not move elements of weight spaces out of their weight space. The action of $N(T)$ on elements of weight spaces descends to an action of $N(T)/T$ on elements of the weight spaces, but in order to know more we would need to calculate $N(T)$.

In SU(2) the copy of T we started with was the one with i in it; that is, we started with $T = \{\left(\begin{smallmatrix} z & 0 \\ 0 & \overline{z} \end{smallmatrix}\right)\} \subset \mathrm{SU}(2)$ (note $\overline{z} = z^{-1}$ since $z \in \mathrm{U}(1)$). We calculate the normalizer of T in SU(2): An element $g \in N(T) \subset \mathrm{SU}(2)$ will normalize all elements $\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right) \in T$; that is, for each $\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right) \in T$ there exists $\left(\begin{smallmatrix} s & 0 \\ 0 & \overline{s} \end{smallmatrix}\right) \in T$ for which $g\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right)g^{-1} = \left(\begin{smallmatrix} s & 0 \\ 0 & \overline{s} \end{smallmatrix}\right)$. Put $v_1 = g^{-1}e_1$, so that

$$g\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right)v_1 = \left(\begin{smallmatrix} s & 0 \\ 0 & \overline{s} \end{smallmatrix}\right)gg^{-1}e_1 = se_1$$

Then $\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right)v_1 = sv_1$, which in other words is to say v_1 is an eigenvector of $\left(\begin{smallmatrix} t & 0 \\ 0 & \overline{t} \end{smallmatrix}\right)$ with eigenvalue s. The eigenvalue s must be equal to one of t or $\overline{t}$. In the case that $s = t$, the element g must be of the form $\left(\begin{smallmatrix} z & 0 \\ 0 & \overline{z} \end{smallmatrix}\right)$ for $z \in \mathrm{U}(1)$. In the case that $s = \overline{t}$, the element g must be of the form $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} z & 0 \\ 0 & \overline{z} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & \overline{z} \\ -z & 0 \end{smallmatrix}\right)$ for $z \in \mathrm{U}(1)$. We can obtain the outcome of either case by taking $t = i$ and solving a small system of equations obtained after conjugation by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in N(T)$ in each case.

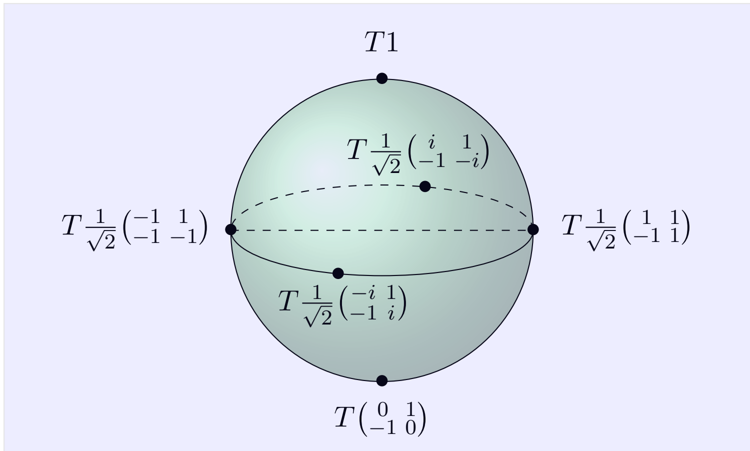
From the above calculations see that $N(T)$ is the group generated by the order 4 element $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and T. The quotient $N(T)/T$ is isomorphic to $\mathbb{Z}/2\mathbb{Z}$ since the coset $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)T$ in $N(T)$ has order 2. Therefore there is a short exact sequence

$$1 \to T \to N(T) \to \mathbb{Z}/2\mathbb{Z} \to 0$$

but this short exact sequence does not split. There is no homomorphism $\mathbb{Z}/2\mathbb{Z} \to N(T)$ which when followed by the quotient map is the identity on $\mathbb{Z}/2\mathbb{Z}$. The generator -1 of $\mathbb{Z}/2\mathbb{Z}$ would need to be sent to an order 2 element of $N(T)$, but $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ has order 4, which is why no section $\mathbb{Z}/2\mathbb{Z} \to N(T)$ exists. So $N(T)$ is a non-split extension of $\mathbb{Z}/2\mathbb{Z}$ by T. As a side remark, if we consider a maximal torus $\widetilde{T}$ inside $\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}_3(\mathbb{R})$, then the normalizer $N(\widetilde{T})$ in $\mathrm{SO}_3(\mathbb{R})$ is a split extension of $\mathbb{Z}/2\mathbb{Z}$ by $\widetilde{T}$. A choice of torus in this case could be, for example, the group of rotations around a fixed axis.

Interestingly, the quotient map $\mathrm{SU}(2) \to T\backslash \mathrm{SU}(2)$ viewed as a quotient map of topological spaces is the Hopf fibration $S^3 \to S^2$, where $T\backslash \mathrm{SU}(2) \cong S^2$. Consider the map $\mathrm{SU}(2) \to \mathbb{C}^2 \setminus \{0\}$ given by $\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right) \to (z, w)$. Then project onto $\mathbb{CP}^1$ by $(z, w) \mapsto [z : w]$. The fibers of the composite map $\mathrm{SU}(2) \to \mathbb{CP}^1$ are the right cosets $T\left(\begin{smallmatrix} z & w \\ -\overline{w} & \overline{z} \end{smallmatrix}\right)$. Therefore we can identify $\mathbb{CP}^1$ with the right coset space $T\backslash \mathrm{SU}(2)$, and $\mathbb{CP}^1 \cong S^2$.

The group $N(T)$ acts on SU(2) by conjugation, and the action descends to an action $N(T)$ on $T\backslash \mathrm{SU}(2)$ by conjugation because $N(T)$ normalizes T. The action of conjugation by an element of T fixes orbits in $T\backslash \mathrm{SU}(2)$, so we can further define an action of $T\backslash N(T) = N(T)/T$ on $T\backslash \mathrm{SU}(T)$ by conjugation. Identifying $T\backslash \mathrm{SU}(2)$ with S^2, the action of $T\backslash N(T)$ looks like reflection across a plane:



In the picture the reflection is across the light blue plane splitting the sphere in half through the middle (through the dashed diameter). Observe that the four points on the top, left, bottom, and right are fixed by the reflection, but the other two points are interchanged. In general, $T\left(\begin{smallmatrix} 0 & t\\ -\bar{t} & 0\end{smallmatrix}\right) \in T\backslash N(T)$ acts by conjugation on an element $T\left(\begin{smallmatrix} z & w\\ -\overline{w} & \overline{z}\end{smallmatrix}\right) \in T\backslash \mathrm{SU}(2)$ by

$$\left(\begin{smallmatrix} 0 & t\\ -\bar{t} & 0\end{smallmatrix}\right) T\left(\begin{smallmatrix} z & w\\ -\overline{w} & \overline{z}\end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -t\\ \bar{t} & 0\end{smallmatrix}\right) = T\left(\begin{smallmatrix} 0 & t\\ -\bar{t} & 0\end{smallmatrix}\right)\left(\begin{smallmatrix} z & w\\ -\overline{w} & \overline{v}\end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -t\\ \bar{t} & 0\end{smallmatrix}\right) = T\left(\begin{smallmatrix} \overline{z} & \overline{w}\\ -w & z\end{smallmatrix}\right).$$

The corresponding action on $\mathbb{CP}^1$ is the involution $[z : w] \mapsto [\overline{z} : \overline{w}]$.

The quotient group $N(T)/T$ is called the Weyl group, denoted W, which in this setting is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. We can also think of the Weyl group as the symmetric group on two elements. The action of W on T by conjugation is to swap the eigenvalues of elements of T, but there is another manifestation of this idea if we return to representations of SU(2).

Let V be a representation of SU(2) and consider its weight space decomposition $V = \bigoplus_{n\in\mathbb{Z}} V_n$ with respect to T. Since T centralizes T, the conjugation action of $N(T)$ on T can be replaced by the faithful conjugation action of $N(T)/T$ on T. We saw before that for $h \in N(T)/T$, $\chi_n(h^{-1}(-)h)$ defines a character of the torus denoted χ_m. Since the action of h by conjugation is by complex conjugation, $\chi_m = \chi_{-n}$. Therefore the action of h on V interchanges V_n with V_{-n} (which also implies $V_n \cong V_{-n}$) for all n.

It follows that characters of representations of SU(2) viewed as polynomials in z are palindromic; that is, if cz^n appears as a term in $\chi_V(z)$, then cz^{-n} must also appear as a term. In particular, for even irreps of SU(2), the only powers of z that could appear in $\chi_V(z)$ are z^n, z^{-n} for n even. An example of a palindromic polynomial that might occur for an irrep V (even or odd, this would just determine if n was even or odd in the following expression) was

$$z^n + 72z^{n-2} + 46z^{n-4} + 46z^{-n+4} + 72z^{-n+2} + z^{-n}$$

We will see much later that this Laurent polynomial actually could not be a character of an irrep V.

For $V = \mathbb{C}^2$ with SU(2) acting on V by matrix multiplication, $\mathbb{C}^2$ has weight space decomposition $\mathbb{C}^2 = \mathbb{C}e_1 + \mathbb{C}e_2$ that is just the standard basis decomposition. The action of an element of the torus T on the first summand is to scale by one eigenvalue (its top-left entry), and the action of an element of the torus T on the second summand is to scale by the other eigenvalue. Therefore the character of this representation is $\chi_V(z) = z + z^{-1}$.

Choice of torus?

At the beginning we chose T to be the maximal torus generated by $1, i$; that is, by the matrices $\left(\begin{smallmatrix} z & 0 \\ 0 & \overline{z} \end{smallmatrix}\right)$ for $z \in \mathrm{U}(1)$. But there are several (infinitely many) other isomorphic copies of U(1) in SU(2) that could have been chosen in place of T. One might object and say that the above theory might have depended on this particular choice of T. We will see that it did not matter.

The first fact we need is that any two maximal toruses in SU(2) are conjugate to each other. This does not constitute a proof, but we can see this visually: A maximal torus of SU(2) is the same as the group of counterclockwise rotations of S^2 about a fixed axis (this picture also exhibits the fact that a maximal torus is isomorphic to $\mathrm{U}(1)^k$ if and only if $k = 1$). Any two maximal toruses then differ only by which axis we choose to rotate about. Obtain an isomorphism of any two maximal toruses by conjugation by a suitable element of SU(2) which rotates one axis to the other. In other words, the action of SU(2) on the set of maximal tori by conjugation is transitive.

The SU(2)-set of maximal tori (with the conjugation action) is isomorphic to the coset space $\mathrm{SU}(2)/\operatorname{Stab}(T) \cong \mathrm{SU}(2)/N(T) \cong (T\backslash \mathrm{SU}(2))/(T\backslash N(T)) \cong S^2/(\mathbb{Z}/2\mathbb{Z})$. The action of $\mathbb{Z}/2\mathbb{Z}$ on S^2 in this case is the antipodal map, so the set of maximal tori can be identified with $\mathbb{RP}^2$. Recall the action of SU(2) on pure quaternions by conjugation; we identified this with the representation $\mathbb{R}^3$ of SU(2). From this representation we obtained a two-to-one map $\mathrm{SU}(2) \to \mathrm{SO}_3(\mathbb{R})$. The action of $\mathrm{SO}_3(\mathbb{R})$ on $\mathbb{R}^3$ is transitive, and therefore the action of SU(2) on $\mathbb{R}^3$ is also. Furthermore, the actions of $\mathrm{SU}(2), \mathrm{SO}_3(\mathbb{R})$ on $\mathbb{R}^3$ descend to transitive actions on $\mathbb{RP}^2$, since the actions move one-dimensional subspaces to one-dimensional subspaces. The stabilizer in $\mathrm{SO}_3(\mathbb{R})$ of a point in $\mathbb{RP}^2$, or rather of a line passing through the origin in $\mathbb{R}^3$, is a copy of $\mathrm{O}_2(\mathbb{R})$ in $\mathrm{SO}_3(\mathbb{R})$. For example, the stabilizer of the line through the standard basis vector e_1 is the copy of $\mathrm{O}_2(\mathbb{R})$ generated by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SO}_2(\mathbb{R})$. So $\mathrm{SO}_3(\mathbb{R})/\mathrm{O}_2(\mathbb{R})$ is isomorphic to $\mathbb{RP}^2$. The preimage of this copy of $\mathrm{O}_2(\mathbb{R})$ in SU(2) is $N(T)$, since the map $\mathrm{SU}(2) \to \mathrm{SO}_3(\mathbb{R})$ is given by

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \mapsto \begin{pmatrix} |z|^2 - |w|^2 & iw\overline{z} - iz\overline{w} & w\overline{z} + z\overline{w} \\ -izw + i\overline{zw} & (w^2 + \overline{w}^2 + z^2 + \overline{z}^2)/2 & -i(w^2 - \overline{w}^2 - z^2 + \overline{z}^2)/2 \\ -zw - \overline{zw} & -i(w^2 - \overline{w}^2 + z^2 - \overline{z}^2)/2 & -(w^2 + \overline{w}^2 - z^2 - \overline{z}^2)/2 \end{pmatrix}$$

The representation $\mathrm{SU}(2) \to \mathrm{SO}_3(\mathbb{R})$ above is sometimes called the adjoint representation of SU(2).

Maximal tori are similar to Sylow subgroups in that the maximal tori are all conjugate to each other and that every element of SU(2) belongs to a maximal torus since conjugation by an element of SU(2) fixes an axis in $\mathbb{R}^3$; that is, conjugation fixes some pure quaternion.

Since any element $g \in \mathrm{SU}(2)$ belongs to a maximal torus T', we may conjugate g by a suitable element $h \in \mathrm{SU}(2)$ so that $hgh^{-1} \in T$. Then $\chi_V(g) = \chi_V(hgh^{-1})$ can be explicitly computed using the Laurent polynomial expression for χ_V obtained from the weight space decomposition of V with respect to T. This shows that the choice of the torus T does not change the form of the Laurent polynomial representing $\chi_V(z)$ (any two such Laurent polynomials would agree on an infinite number of points).

Matrix elements

We take a detour to discuss matrix elements, which will explain where characters come from, and will lead to the Peter-Weyl theorem.

From $v \in V$ and $w \in V^*$, define $f_{v,w} \colon G \to \mathbb{C}$ by $g \mapsto \langle w, gv\rangle := w(gv)$. We call the functions $f_{v,w}$ for $v \in V$ and $w \in V^*$ matrix elements, and $f_{v,w}(g)$ the matrix elements of g. Note that this definition in terms of V and V^* is coordinate-free, but recovers what we think are "matrix elements". If we fix a basis $\{e_i\}$ of V and give V^* the dual basis, then $f_{e_j^*,e_k}(g) = \langle e_j^*, ge_k\rangle$ denotes the (j,k)-entry in the matrix for $g\cdot$ in the given basis for V. Collect all the matrix elements via a map $V \otimes V^* \xrightarrow{f_{-,-}} \mathrm{Fun}(G)$ (defined on simple tensors, extended linearly).

A different perspective: Let V be a finite-dimensional representation of a group G; that is, consider a map $G \xrightarrow{\rho} \mathrm{GL}(V) \subset \mathrm{End}(V) \cong V \otimes V^*$. Given any map F in $(V \otimes V^*)^*$ we may pull back F along ρ to obtain a map $G \xrightarrow{\rho^* F} \mathbb{C}$. So ρ^* defines a map from $(V \otimes V^*)^*$ to $\mathrm{Fun}(G)$. The identification $(V \otimes V^*)^* \cong V^* \otimes V^{**} \cong V^* \otimes V \cong V \otimes V^*$ shows that the matrix elements map $f_{-,-}$ agrees with ρ^*, which is really some kind of adjoint map (i.e., an adjoint map of vector spaces) to the map ρ from G (or by linearly extending ρ, from a suitably chosen $\mathrm{Fun}(G)$) to $\mathrm{End}(V) \cong V \otimes V^*$.

We would like to investigate what kinds of functions the matrix elements are. For example, if V is a continuous representation, then the matrix elements are continuous functions, and if G is compact, then the matrix elements are also L^2-integrable.

Since V is a representation of G, the vector space $V \otimes V^*$ is a $G \times G$-representation by the action $(g,h)(v \otimes w) = hv \otimes gw$. The reason this looks backwards is because we identified $V^* \otimes V$ with $V \otimes V^*$ by the map $w \otimes v \to v \otimes w$; really the matrix elements map should have been defined on $V^* \otimes V$ to begin with. Furthermore, we can identify $V \otimes V^*$ with $\mathrm{End}(V)$ as vector spaces by the map $v \otimes w \mapsto w(-)v$, but we can upgrade the isomorphism to an isomorphism of $G \times G$-representations if we let $G \times G$ act on $\mathrm{End}(V)$ by $(g,h)f = hf(g^{-1}(-))$. The space $\mathrm{Fun}(G)$ is also a $G \times G$-representation where $(g,h)f = f(g^{-1}(-)h)$ (combining both the left and right actions of G on

Fun(G)). Then the matrix elements map $V \otimes V^* \to \mathrm{Fun}(G)$ is a $G \times G$-equivariant map:

$$hv \otimes gw \mapsto f_{hv,gw} = \langle gw, (-)hv\rangle = \langle w, g^{-1}(-)hv\rangle = gf_{v,w}h$$

Let $n = \dim V$. By forgetting the action of G on V^*, what was the $G \times G$-representation $V \otimes V^*$ may be thought of as a representation of G whose action is concentrated in V (that is, G acts trivially on V^*). Then as G-representations, $V \otimes V^*$ is isomorphic to the n-fold direct sum of copies of V. So by forgetting the left action of G on Fun(G) (that is, to consider Fun(G) as a G-representation by the action $hf = f((-)h)$), the matrix elements map is a G-equivariant map $V \otimes V \xrightarrow{f_{-,-}} \mathrm{Fun}(G)$.

Now assume V is an irrep of G. Then $V \otimes V^*$ is an irrep of $G \times G$ (See this MSE post). In this case the matrix element map $f_{-,-}$ is either zero or injective (maps out of simple modules are either zero or injective). If V is not the zero space, then $V \otimes V^* \cong \mathrm{End}(V)$ is not trivial since it has id_V. To find id_V in $V \otimes V^*$, we fix a basis $\{e_i\}$ of V and give V^* the corresponding dual basis. Then id_V corresponds to the element $\sum_i e_i \otimes e_i^*$. The element $\mathrm{id}_V \in V \otimes V^*$ is sent to the matrix element $g \mapsto \langle e_i^*, ge_i\rangle = \mathrm{Tr}_V(g\cdot)$, which is the character χ_V (of course, this means the whole subspace $\mathbb{C}\,\mathrm{id}_V \subset V \otimes V^*$ is sent to the subspace $\mathbb{C}\chi_V \subset \mathrm{Fun}(G)$). In particular, notice that $\chi_g(\mathrm{id}_G) = \dim V \neq 0$. Therefore the matrix elements map $f_{-,-}$ is a nonzero map and hence is an injective $G \times G$-equivariant map in the case that V is an irrep.

Like before, by forgetting the actions of one of the factors of G on $V \otimes V^*$ and Fun(G), we still obtain an injective map of G-representations. So in particular for V an irrep, the matrix elements map $f_{-,-}$ manifests $n = \dim V$ copies of the irrep V in Fun(G). This is to say that there are (at least) n copies of V in Fun(G) since the matrix elements map is injective. In the case that G is a finite group, we can recover the formula that the regular representation $\mathrm{Fun}(G) \cong \mathbb{C}[G] \cong \bigoplus_{V \text{ irrep}} V^{\dim V}$ from finite group representation theory, which is a kind of Peter-Weyl theorem, as we will see.

Since V is an irrep, Schur's lemma says that the space $\mathrm{End}_G(V)$ of G-equivariant endomorphisms of V is equal to $\mathbb{C}\,\mathrm{id}_V$. Consider the diagonal embedding of G into $G \times G$ by $g \mapsto (g, g)$. Through this map we can think of $V \otimes V^*$ as a G-representation by the action $g(v \otimes w) = (g, g)(v \otimes w) = gv \otimes gw$, and similarly of Fun($G$) by $gf = (g, g)f = f(g^{-1}(-)g)$. Observe that the subspace $(V \otimes V^*)^G$ of G-invariants in $V \otimes V$ is isomorphic to $\mathrm{End}_G(V)$, which is isomorphic to $\mathbb{C}\,\mathrm{id}_V$ since V is an irrep. Therefore the image of $(V \otimes V^*)^G \cong \mathbb{C}\,\mathrm{id}_V$ under $f_{-,-}$, $\mathbb{C}\chi_V$, is contained in the space $\mathrm{Fun}(G)^G$ of (continuous/integrable/etc.) class functions as well, since intertwining maps preserve G-invariance. This is another way to see that the character χ_V is invariant under conjugation. If V is not an irrep, then we can still see that the image of $(V \otimes V^*)^G \cong \mathrm{End}_G(V)$ under $f_{-,-}$, whatever it may be, is still contained in $\mathrm{Fun}(G)^G$.

There may be some confusion about why we call χ_V a character if it is not quite the same thing as the characters we looked at before in the Abelian group case. If G is compact Abelian, then the character $\chi_V = \mathrm{Tr}_V$ of an irrep V of G will agree with whatever character χ the action of G on V is given by. That is, since V is unitarizable and one-dimensional, G will act on V by multiplication by $\chi(g)$ for some character χ, and the trace of $g\cdot$ is exactly $\chi(g)$ as a result. In general for non-Abelian groups we should not expect to match characters of representations with characters of their groups, but the previous example might explain why the name "character" is used for both. As an example, there are no nontrivial group homomorp hisms $\mathrm{SU}(2) \to \mathrm{U}(1)$, but we can and should consider characters of representations of $\mathrm{SU}(2)$.

Preview of the Peter-Weyl theorem

Let V, W be irreps of a compact group G. The matrix elements map on $(V \oplus W) \otimes (V \oplus W)^*$ is $(v, w) \otimes (s, t) \mapsto f_{(v,w),(s,t)}$ where

$$f_{(v,w),(s,t)}(g) = \langle (s,t), (gv, gw) \rangle = \langle (s,0), (gv,0) \rangle + \langle (s,0), (0,gw) \rangle + \langle (0,t), (gv,0) \rangle + \langle (0,t), (0,gw) \rangle$$

The middle two terms are zero, and using the identification $(V \oplus W) \otimes (V \oplus W)^* \cong (V \otimes V^*) \oplus (W \otimes V^*) \oplus (V \otimes W^*) \oplus (W \otimes W^*)$, we see that $f_{(v,w),(s,t)} = f_{v,s} + f_{w,t}$, where the $f_{-,-}$ appearing on the right side are on $V \otimes V^*$ and $W \otimes W^*$, respectively. So the matrix elements map on $(V \oplus W) \otimes (V \oplus W)^*$ really restricts to the direct sum of the matrix elements maps $f_{-,-} \oplus f_{-,-}$ on $(V \otimes V^*) \oplus (W \otimes W^*) \cong \mathrm{End}(V) \oplus \mathrm{End}(W)$. Since V, W are irreps, $f_{-,-}$ on each of $(V \otimes V^*)$ and $(W \otimes W^*)$ is injective, so the direct sum $f_{-,-} \oplus f_{-,-}$ is injective and hence the matrix elements map $f_{-,-}$ on $(V \oplus W) \otimes (V \oplus W)^*$ is also injective. We did not need to assume that W was distinct from V since the action of G preserves each summand in $V \oplus W$.

We can be ambitious and try to move from two distinct irreps to every irrep of G counted once. That is, we look at the matrix elements map on $\bigoplus_{V \text{ irrep}} V \otimes V^*$, and the same argument as above shows that the matrix elements map defined on this direct sum is also injective, mapping into $\mathrm{Fun}(G)$, or really into $C(G)$ or $L^2(G)$. The image of the subspace $\bigoplus_{V \text{ irrep}} \mathbb{C}\,\mathrm{id}_V$ in $L^2(G)$ is $\bigoplus_{V \text{ irrep}} \mathbb{C}\chi_V$. We can give $L^2(G)$ a nice inner product $\langle -, - \rangle$ given by $\langle f, g \rangle = \int_G f\overline{g}\mathrm{d}\mu$ where μ is the normalized Haar measure for G; by doing so the characters χ_V for irreps V form an orthogonal family in $L^2(G)$. Note that the Haar measure μ is $G \times G$-invariant (left and right-translation invariant), so the inner product on $L^2(G)$ is also $G \times G$-invariant as well.

Part of the statement of the Peter-Weyl theorem is as follows: The image of $\bigoplus_{V \text{ irrep}} V \otimes V^*$ in $L^2(G)$ under the matrix elements map is dense in $L^2(G)$. If G is $\mathrm{U}(1)$, then the Peter-Weyl theorem states that the Fourier modes $\exp(inx)$ form a dense orthogonal basis of $L^2(\mathbb{R}/\mathbb{Z})$.

The functions coming from the image of $\bigoplus_{V \text{ irrep}} V \oplus V^*$ in $L^2(G)$ are the algebraic functions on G; that is, we can identify $\bigoplus_{V \text{ irrep}} V \oplus V^*$ with $\mathbb{C}[G]$ (where this notation means the coordinate ring of G, and since the matrix elements map is injective we might as well just think about $\bigoplus_{V \text{ irrep}} V \oplus V^*$ instead of its image). In the case $G = \mathrm{U}(1)$, $\mathbb{C}[G]$ is $\mathbb{C}[z, z^{-1}]$.

Lecture 10 September 25

The Peter-Weyl theorem

Let G be a compact Lie group and let V be a finite-dimensional (continuous, we have been and will continue to assume this) representation of G. Since V is a continuous representation of G, the matrix elements map $f_{-,-}$ defined last time has image lying in $C(G)$; that is, the matrix elements $f_{v,w}$ for $v \in V$ and $w \in V^*$ are continuous functions.

Last time we also saw that the matrix elements map $f_{-,-}$ was a $G \times G$-intertwining map by carefully defining the actions of each factor of G on $V \otimes V^*$ and $\mathrm{Fun}(G)$. Furthermore, we saw that if V was an irrep, then $V \otimes V^*$ was a $G \times G$-irrep so the matrix elements map $f_{-,-}$ was injective. And if V, W were irreps, then the matrix elements map on $(V \oplus W) \otimes (V \oplus W)^*$ decomposed into the injective map $f_{-,-} \oplus f_{-,-}$ on $(V \otimes V^*) \oplus (W \otimes W^*)$, which implies the matrix elements map on $(V \oplus W) \otimes (V \oplus W)^*$ was injective to begin with. Then we defined the matrix elements map on $\bigoplus_{V \text{ irrep}} V \otimes V^*$, which maps injectively into $C(G)$.

But the expression $\bigoplus_{V \text{ irrep}} V \otimes V^*$ maybe isn't so clear, since we would like for any irrep W of G to find $W \otimes W^*$ in the direct sum $\bigoplus_{V \text{ irrep}} V \otimes V^*$. For each irrep V, fix a basis of V and give V^* the corresponding dual basis. Then by Schur's lemma there is only one $G \times G$-equivariant map of irreps $W \otimes W^* \to V \otimes V^*$ which sends id_W to id_V. Furthermore, id_V and id_W both map to the same function under the matrix elements map, and in general the matrix elements map would agree on both $V \otimes V^*$ and $W \otimes W^*$. So it would not matter which isomorphism class of each irrep V we take in the direct sum $\bigoplus_{V \text{ irrep}} V \otimes V^*$. Henceforth fix an enumeration of the irreps V of G, and refer to these irreps only.

Each irrep V is unitarizable, so fix a G-invariant inner product on each V. The (antilinear) Riesz map $V \xrightarrow{r} V^*$ given by $v \mapsto \langle -, v \rangle$ is G-equivariant:

$$gv \mapsto \langle -, gv \rangle = \langle gg^{-1}(-), gv \rangle = \langle g^{-1}(-), v \rangle = g\langle -, v \rangle$$

Therefore the inverse to the Riesz map is also G-equivariant, so we can give a natural G-invariant inner product on V^*:

$$\langle -, - \rangle_{V^*} = \langle r^{-1}(-), r^{-1}(-) \rangle_V$$

Later we will give $\bigoplus_{V \text{ irrep}} V \otimes V^*$ a $G \times G$-invariant inner product for which the elements id_V in $\bigoplus_{V \text{ irrep}} V \otimes V^*$ are orthogonal. This inner product can also be chosen so that the matrix elements map $f_{-,-}$ is a map of unitary representations where $C(G)$ is given a suitable weighted L^2 inner product. We will see later that this implies the orthogonality of the characters χ_V for irreps V.

We can give $\bigoplus_{V \text{ irrep}} V \otimes V^*$ a (commutative) ring structure that is not the (noncommutative) ring structure coming from function composition in $\bigoplus_{V \text{ irrep}} \mathrm{End}(V)$. This is done by pulling back the commutative ring structure on $\mathrm{Fun}(G)$. For $v \otimes w, u \otimes z \in V \otimes V^*$, since we defined the matrix elements map $f_{-,-}$ on simple tensors and extended by linearity, the addition of $v \otimes w, u \otimes z$ in $\bigoplus_{V \text{ irrep}} V \otimes V^*$ is just the usual addition in the summand $V \otimes V^*$.

Let V, W be irreps of G. For $v \otimes v' \in V \otimes V^*$ and $w \otimes w' \in W \otimes W^*$, the addition of matrix elements is given pointwise by

$$(f_{v,v'} + f_{w,w'})(g) = f_{v,v'}(g) + f_{w,w'}(g) = \langle v', gv\rangle + \langle w', gw\rangle = \langle (v', w'), g(v, w)\rangle = f_{(v,w),(v',w')}(g)$$

and the multiplication of matrix elements is also given pointwise by

$$(f_{v,v'} \cdot f_{w,w'})(g) = f_{v,v'}(g) \cdot f_{w,w'}(g) = \langle v', gv\rangle\langle w', gw\rangle = \langle v' \otimes w', g(v \otimes w)\rangle = f_{v\otimes w, v'\otimes w'}(g)$$

This shows that the ring structure on $\bigoplus_{V \text{ irrep}} V \otimes V^*$ is given by basically taking the direct sum and tensor product, where $(v\otimes v') + (w\otimes w') = (v, w)\otimes(v', w')$ and $(v\otimes v')\cdot(w\otimes w') = (v\otimes w)\otimes(v'\otimes w')$. But the elements $(v, w) \otimes (v', w')$ and $(v \otimes w) \otimes (v' \otimes w')$ are not necessarily elements of $\bigoplus_{V \text{ irrep}} V \otimes V^*$. So far, these addition and multiplication formulas have been formulated so that the matrix elements map yields the same function when applied to both sides of each formula (we don't need to break representations into direct sums of irreps to define the matrix elements map, but for the purposes of defining a ring structure on $\bigoplus_{V \text{ irrep}} V\otimes V^*$ this is worth observing). To define the ring operator properly in $\bigoplus_{V \text{ irrep}} V \otimes V^*$, the addition and multiplication formulas should return elements of that space.

We start with the addition of simple tensors. We had $(v\otimes v') + (w\otimes w') = (v, w)\otimes(v', w') \in (V \oplus W)\otimes(V \oplus W)^*$. Decompose $(V \oplus W)\otimes(V \oplus W)^*$ into $(V \otimes V^*)\oplus(W \otimes V^*)\oplus(V \otimes W^*)\oplus(W \otimes W^*)$, on which we saw earlier that the matrix elements map was zero on the middle two summands, regardless of whether V was distinct from W or not. This decomposition is just $(v, w)\otimes(v', w') = (v, 0)\otimes(v', 0) + (0, w)\otimes(v', 0) + (v, 0)\otimes(0, w') + (0, w)\otimes(0, w')$. Because the matrix elements map is trivial on the middle two terms, we can ignore (delete) these terms. The first and fourth terms can be combined if $V = W$ (remember we have fixed an enumeration of every irrep of G, so equals means equals). If $V = W$, then send $(v, 0) \otimes (v', 0) + (0, w) \otimes (0, w')$ to $v \otimes v' + w \otimes w' \in V \otimes V^*$. Otherwise, view the first and fourth terms in $\bigoplus_{V \text{ irrep}} V \otimes V^*$ as $(v \otimes v') + (w \otimes w')$ (suppressing the notation for elements in a direct sum). To summarize, if V and W are distinct, we have defined a map $(V \oplus W) \otimes (V \oplus W)^* \to (V \otimes V^*) \oplus (W \otimes W^*) \hookrightarrow \bigoplus_{V \text{ irrep}} V \otimes V^*$. In the case that V and W coincide, we have defined a similar map $(V \oplus W) \otimes (V \oplus W)^* \to (V \otimes V^*) \hookrightarrow \bigoplus_{V \text{ irrep}} V \otimes V^*$. Denote both maps by $\underline{\text{add}}$; that is, $\underline{\text{add}}((v, w) \otimes (v', w')) = v \otimes v' + w \otimes w'$ where depending on whether V is distinct from W or not, we view $v \otimes v' + w \otimes w'$ in $(V \otimes V^*) \oplus (W \otimes W^*)$ or $V \otimes V^*$. Then the following diagram commutes:

$$\begin{tikzcd}
\bigoplus_{V \text{ irrep}} V \otimes V^* \arrow[r, "f_{-,-}"] & \mathrm{Fun}(G) \\
(V \oplus W) \otimes (V \oplus W)^* \arrow[u, "\underline{\text{add}}"] \arrow[ur, "f_{-,-}"'] &
\end{tikzcd}$$

So the correct addition law in $\bigoplus_{V \text{ irrep}} V \otimes V^*$ is $(v \otimes v') + (w \otimes w') = \underline{\text{add}}((v, w) \otimes (v', w')) = v \otimes v' + w \otimes w'$, where the notation on the right hand side is confusingly not the same as the notation on the left hand side.

We do a similar thing with multiplication. On simple tensors, we had $(v \otimes v') \cdot (w \otimes w') = (v \otimes w) \otimes (v' \otimes w') \in (V \otimes W) \otimes (V \otimes W)^*$. Decompose $(V \otimes W) \otimes (V \otimes W)^*$ into

$$\left(\bigoplus_i U_i\right) \otimes \left(\bigoplus_j U_j^*\right)$$

where U_i and U_j are irreps of G; more importantly, the irreps appearing the direct sums may repeat. We define a similar map to add which coalesces the elements in repeated summands U_i by addition, which yields an element of $\bigoplus_{V \text{ irrep}} V \otimes V^*$. The coalesce map is defined in the same way as add above, and we can even just use this map in place of add above. We obtain the commutative diagram

$$\begin{tikzcd}
\bigoplus_{V \text{ irrep}} V \otimes V^* \arrow[r, "f_{-,-}"] & \mathrm{Fun}(G) \\
\left(\bigoplus_i U_i\right) \otimes \left(\bigoplus_j U_j^*\right) \arrow[u, "\underline{\text{coalesce}}"] \arrow[ur, "f_{-,-}"'] &
\end{tikzcd}$$

It follows that there is a law for multiplication on $\bigoplus_{V \text{ irrep}} V \otimes V^*$, but the explicit form of this law is unclear since we would need to know how the tensor product of irreps decompose as a direct sum of irreps to carry out the multiplication. In practice this entire technicality should not matter and we would not need to insist on obtaining elements of $\bigoplus_{V \text{ irrep}} V \otimes V^*$ after addition or multiplication.

Call $\bigoplus_{V \text{ irrep}} V \otimes V^*$ the ring of matrix elements of G. By defining the ring of matrix elements this way, the matrix elements map $f_{-,-}$ is a ring homomorphism.

The ring structure on $\bigoplus_{V \text{ irrep}} V \otimes V^* \cong \bigoplus_{V \text{ irrep}} \mathrm{End}(V)$ coming from function composition is also relevant. The function space $C(G)$ with the convolution product is a non-commutative ring (for G non-Abelian). Define the action map $\mathrm{Act}\colon C(G) \to \bigoplus_{V \text{ irrep}} \mathrm{End}(V)$ by

$$h \mapsto \int_G h(g) g(-) \mathrm{d}g$$

(this is the Bochner integral using the normalized Haar measure on G we saw before in an earlier discussion about the spectral theorem on $\mathrm{U}(1)$). So in some sense the function space $C(G)$ is acting as a version of the group algebra, but is not quite there since we are missing notable elements like the delta distribution, which is the unit of the convolution product. Observe that $\mathrm{Act}(\mathbf{1}_G)$ is the averaging operator $\mathrm{av} \in \bigoplus_{V \text{ irrep}} \mathrm{End}(V)$, which defines a projector to the direct sum of trivial representations.

The inner product of the matrix element $f_{v,v'} = \langle v', (-)v \rangle$ with a function h is

$$\langle f_{v,v'}, h \rangle = \int_G \langle v', gv \rangle \overline{h(g)} \mathrm{d}g = \left\langle v', \int_G \overline{h(g)} gv \right\rangle = \langle v', \mathrm{Act}(\overline{h})(v) \rangle$$

It follows that $f_{v,-}$ and $\mathrm{Act}(\overline{-})(v)$ are adjoint to each other in some sense. Compare this with the earlier remark about how the matrix elements map $f_{-,-}$ is a kind of adjoint (pullback) of ρ for a representation (V, ρ) of G.

Let V be a representation of G which we do not assume is finite-dimensional. Then denote by V^{fin} the subspace of V containing vectors which belong to some finite-dimensional subrepresentation of V (we are not fixing a finite-dimensional subrepresentation of V for this definition, we just insist that the vectors live in a finite-dimensional subrepresentation of V). Using this definition, we observe the following: The ring of matrix elements is isomorphic to $C(G)^{\text{fin}}$ as G-representations (or rather the image of the ring of matrix elements under the matrix elements map is equal to $C(G)^{\mathrm{fin}}$), where the action of G on matrix elements is given by $h(v \otimes v') = hv \otimes v'$ and the action of G on $\mathrm{Fun}(G)^{\mathrm{fin}}$ is $hf = f((-)h)$. The matrix elements map is injective, so it suffices to see that the image of the matrix elements map is $C(G)^{\mathrm{fin}}$. Assume that f is in $C(G)^{\mathrm{fin}}$, that is, $f \in W$ for some finite-dimensional subrepresentation W of $C(G)$. Form a basis $\{f_i\}$ of W where $f_1 = f$. Let $\{f_i^*\}$ be the corresponding dual basis for W^*. Then the action of h on each f_i is given by

$$hf_i = \sum_j \langle f_j^*, hf_i \rangle f_j$$

From this equation and $f_1 = f$ we obtain

$$f(h) = (hf_1)(1) = \sum_j \langle f_j^*, hf_1 \rangle f_j(1) = \sum_j f_j(1) f_{f_1, f_j^*}(h)$$

It follows that any function in $C(G)^{\mathrm{fin}}$ can be obtained by a linear combination of matrix elements, which proves the proposition. As an example, if $G = \mathrm{U}(1)$, then $C(G)^{\mathrm{fin}}$ is the space of functions given by finite Fourier series.

Since $C(G)$ is an inner product space with the usual L^2 inner product, we may pull back this inner product to obtain some inner product on $\bigoplus_{V \text{ irrep}} V \otimes V^*$, which is a direct sum of inner products on each irrep $V \otimes V^*$, because a G-invariant inner product on $(V \otimes V^*) \oplus (W \otimes W^*)$ for distinct irreps V, W of G is an antilinear G-equivariant isomorphism of $(V \otimes V^*) \oplus (W \otimes W^*)$ with its dual. But by Schur's lemma such an isomorphism is really a direct sum of antilinear G-equivariant isomorphisms of $(V \otimes V^*)$ with its dual and $(W \otimes W^*)$ with its dual. The inner product on $V \otimes V^*$ is given on simple tensors by

$$\langle v \otimes v', w \otimes w' \rangle_{V \otimes V^*} = C_V \langle v, w \rangle_V \langle v', w' \rangle_{V^*}$$

where C_V is a constant depending on V. Let $\{e_i\}$ be a basis of V and give V^* the corresponding dual basis. Then to calculate C_V, calculate the L^2 inner product of the matrix elements f_{e_i, e_j^*} and f_{e_s, e_t^*}. The inner product on $V \otimes V^*$ is obtained by pulling back this inner product, so

$$C_V \langle e_i, e_s \rangle_V \langle e_j^*, e_t^* \rangle_{V^*} = \langle f_{e_i, e_j^*}, f_{e_s, e_t^*} \rangle_{L^2} = \int_G f_{e_i, e_j^*}(g) \overline{f_{e_s, e_t^*}(g)} \mathrm{d}g = \int_G M_{ij}(g) \overline{M_{st}(g)} \mathrm{d}g$$

where $M_{kl}(g)$ is the (k, l)-entry of the unitary matrix representing the action of g on V. Since g acts by a unitary matrix, $\overline{M_{st}(g)} = M_{ts}(g^{-1})$. Observe further that $\langle e_i, e_s \rangle = \delta_{is}$ and $\langle e_j^*, e_t^* \rangle_{V^*} = \langle e_j, e_t \rangle_V = \delta_{jt}$. Then

$$C_V \delta_{is} \delta_{jt} = \int_G M_{ts}(g^{-1}) M_{ij}(g) \mathrm{d}g,$$

and by setting $i = s$ and summing over i, obtain

$$C_V \delta_{jt} \dim V = \int_G \sum_{i=1}^{\dim V} M_{ti}(g^{-1}) M_{ij}(g) \mathrm{d}g = \int_G M_{tj}(g) \mathrm{d}g = \delta_{tj}$$

Hence $C_V = 1/\dim V$. So the inner product on $\bigoplus_{V \text{ irrep}} V \otimes V^*$ is given by the direct sum of the inner products$(1/\dim V) \langle -, - \rangle_V \langle -, - \rangle_{V^*}$.

This immediately implies the orthonormality of the characters χ_V for irreps V, since they are obtained as the matrix elements of the identity maps $\mathrm{id}_V = \sum_i e_i \otimes e_i^*$ in each $V \oplus V^*$. Therefore any finite dimensional representation of G is determined up to isomorphism by its character. We observed earlier that the characters χ_V actually belong to $C(G)^G$, the space of class functions (here $C(G)$ is a G-representation by the diagonal embedding $G \to G \times G$ which results in the conjugation action, and we may further observe that the characters belong to $(C(G)^{\mathrm{fin}})^G$). We will see that $(C(G)^{\mathrm{fin}})^G$ is L^2-dense in $C(G)^G$ and hence also $L^2(G)^G$. Observe that the characters from an orthonormal basis of $C(G)^G$ (or of $L^2(G)^G$), since the G-invariant (conjugation-invariant) subspace of $\bigoplus_{V \text{ irrep}} V \otimes V^*$ is $\bigoplus_{V \text{ irrep}} \mathbb{C}\,\mathrm{id}_V$. We will see by analysis that $C(G)^{\mathrm{fin}}$ is L^2-dense in $C(G)$ and hence also $L^2(G)$.

If U, P are any representations of G, then by decomposing U and P into irreps (which may repeat), observe that $\langle \chi_U, \chi_P \rangle_{L^2} = \dim \mathrm{Hom}_G(U, P)$. The dimension of $\mathrm{Hom}_G(U, P)$ simply counts the number of ways irreps in U map to irreps in P, which is also what the inner product of the characters is counting, by the orthonormality of characters of irreps.

It turns out that Act defines an isomorphism of algebras $C(G)^{\mathrm{fin}}$ with convolution and $\bigoplus_{V \text{ irrep}} V \otimes V^* \cong \bigoplus_{V \text{ irrep}} \mathrm{End}(V)$ with composition. Finally, it also turns out that the polynomial functions on G are the same as $C(G)^{\mathrm{fin}}$, which will connect these ideas to algebraic geometry.

Return to SU(2) and Fourier theory

Observe that SU(2) lives inside the space $\mathrm{End}(\mathbb{C}^2)$ of 2×2-complex valued matrices; in other words, there is a faithful two-dimensional representation of SU(2) (faithful meaning $\mathrm{SU}(2) \to \mathrm{GL}(V)$ is injective). We can look at the polynomial functions on SU(2), and by Stone-Weierstrass (using the identification $\mathrm{End}(\mathbb{C}^2) \cong \mathbb{R}^8$ as real vector spaces) these polynomial functions are dense in $C(\mathrm{SU}(2))$, and hence also in $L^2(\mathrm{SU}(2))$. Since the polynomial functions on SU(2) are $C(\mathrm{SU}(2))^{\mathrm{fin}}$, that is, they are the matrix elements of SU(2), then we can recover the Peter-Weyl theorem directly without using the theory from above.

A related corollary of the Peter-Weyl theorem is that every compact Lie group is isomorphic to a compact subgroup of $\mathrm{GL}_n(C)$; that is, every compact Lie group has a faithful finite-dimensional representation.

Observe further that if V is any (unitarizable) representation of SU(2) or G a compact Lie group, then V^{fin} is dense in V (as a Hilbert space). It suffices to show that $C(G)^{\mathrm{fin}}$ is dense in $C(G)$. Recall that $C(G)$ acts on $C(G)$ by the action map Act, by $f \mapsto \int_G f(g)g(-)\mathrm{d}g$; that is, $C(G)$ acts on itself by convolution. One can check that the endomorphism $\mathrm{Act}(f)$ is a compact operator and that if $\overline{f(g)} = f(g^{-1})$, then $\mathrm{Act}(f)$ is self-adjoint. An example of such an f for which $\mathrm{Act}(f)$ is self-adjoint is χ_V for V a unitarizable representation of G. Since g acts as a unitary operator U on V, g^{-1} acts by $\overline{U}^\top$, so $\overline{\chi_V(g)} = \chi_V(g^{-1})$. Self-adjoint compact operators have finite-dimensional eigenspaces and any function in the range of such an operator may be uniformly approximated by eigenfunctions. Take f to be a class function so that the action of G on $C(G)$ commutes with $\mathrm{Act}(f)$; therefore eigenfunctions of $\mathrm{Act}(f)$ belong to $C(G)^{\mathrm{fin}}$. Since there is a sequence f_n of class functions approximating the delta distribution δ_{1_G}

(which acts on $C(G)$ by the identity), it follows that $C(G)^{\mathrm{fin}}$ is dense in $C(G)$ and hence also $L^2(G)$. By restricting to eigenfunctions which are class functions we obtain further that $C(G)^G$ is dense in $L^2(G)^G$. That V^{fin} is dense in V for any representation V of G also implies that every irrep of a compact Lie group is finite-dimensional.

Since $L^2(G)^G$ is isomorphic to $\widehat{\bigoplus}_{V \text{ irrep}}(V \otimes V^*)^G \cong \widehat{\bigoplus}_{V \text{ irrep}}\mathbb{C}\chi_V$ as Hilbert spaces, we can define $\widehat{G}$ to be the collection of irreducible unitary representations of G, and define the sequence space $\ell^2(\widehat{G})$ with the measure $1/\dim$, which is the counting measure but weighted so that for an irrep V, the set $\{V\}$ has measure $1/\dim V$. So we can see that $L^2(G)^G$ is isomorphic to $\ell^2(\widehat{G})$ as Hilbert spaces. The Fourier series in this case is to expand as a series of characters χ_V.

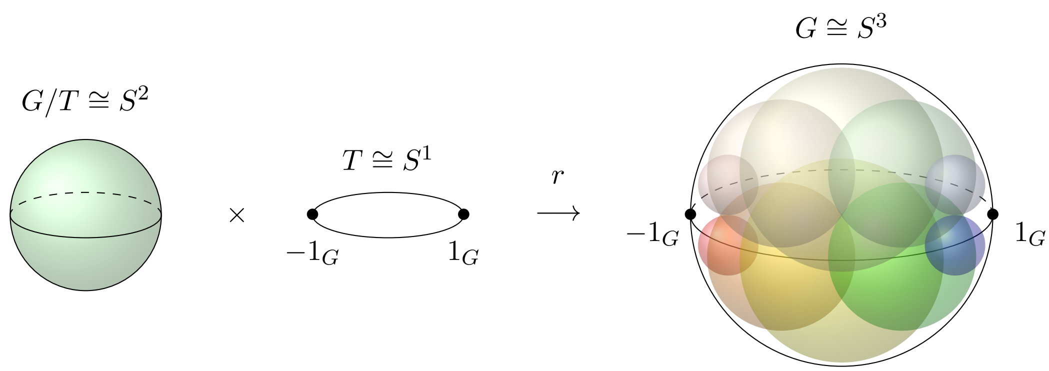
Lecture 11 September 30

Weyl integration and character formulas

We know from the Peter-Weyl theorem that $L^2(G)^G$, the L^2-integrable class functions on G, is isomorphic to $\ell^2(\widehat{G}, 1/\dim)$ where $\widehat{G}$ is the set of irreducible unitary representations of G; an orthonormal basis of $L^2(G)^G$ is given by characters of irreps of G.

Let $G = \mathrm{SU}(2)$. We can restrict any class function on G to a maximal torus T of G (any maximal torus will do since they are all conjugate to each other) to obtain a palindromic function on T, since we saw before that characters of representations restrict to palindromic polynomials on T. Denote the subspace of $L^2(T)$ of palindromic functions by $L^2(T)^W$. These palindromic functions should be thought of as even functions on a circle, since they are invariant under conjugation. But the restriction map $L^2(G)^G \to L^2(T)^W$ is not an isometry. The Weyl integration formula measures the difference.

Consider the following "radial parts construction". The surjective map $G \times T \to G$ given by $(g, t) \mapsto gtg^{-1}$ descends to a surjective map $G/T \times T \xrightarrow{r} G$ given by $(\overline{g}, t) \mapsto gtg^{-1}$ since T centralizes T. The map r is a local diffeomorphism and has degree two; indeed, the preimage of a generic point gtg^{-1} is $\{(\overline{g}, t), (\overline{gw}, w^{-1}tw)\}$, where $\overline{w} \in N(T)/T \cong \mathbb{Z}/2\mathbb{Z}$ is the nonidentity element. So $G/T \times T$ is a two-fold cover of G. The map r kind of looks like



Observe that since T has $\pm 1_G$, the spheres at the right and left sides of the S^3 above picture really are pinched to a point (the identity conjugates to the identity).

We will take for granted that G/T has a left and right G-invariant (bi-invariant) measure normalized so that G/T, which is compact, has measure 1. This comes from the bi-invariant measure on G. Taking the product with the Haar measure on T, form a measure μ on $G/T \times T$. If f is a function on G, then

$$\int_G f \mathrm{d}g = \frac{1}{2} \int_{G/T \times T} r^*(f \mathrm{d}g) = \frac{1}{2} \int_{G/T \times T} r^* f r^*(\mathrm{d}g) = \frac{1}{2} \int_{G/T \times T} (r^* f) J \mathrm{d}\mu$$

where J is the Jacobian of r. The appearance of the factor $1/2$ after the first equality is because r is generically a two-to-one map. Now assume f is a class function, so $(r^*f)(\overline{g},t) = f(gtg^{-1}) = f(t)$. Then

$$\frac{1}{2}\int_{G/T\times T}(r^*f)J\mathrm{d}\mu = \frac{1}{2}\int_T f\int_{G/T} J\mathrm{d}t\mathrm{d}\overline{g}$$

Because the measure on G/T is bi-invariant, it turns out that J does not depend on G/T. That is, J is a function of T only, so

$$\frac{1}{2}\int_T f\int_{G/T} J\mathrm{d}t\mathrm{d}\overline{g} = \frac{1}{2}\int_T fJ\int_{G/T}\mathrm{d}\overline{g}\mathrm{d}t = \frac{1}{2}\int_T fJ\mathrm{d}t$$

We calculate J. The Lie algebra $\mathfrak{t}$ of T lives inside the Lie algebra $\mathfrak{g}$ of G, and we can identify the tangent space of G/T at $\overline{1}_G$ with the orthogonal complement of $\mathfrak{t}$ in $\mathfrak{g}$ (with respect to the Killing form on $\mathfrak{g}$). A back-of-the-envelope calculation shows that by perturbing $t\in T$ in the direction of $\xi\in\mathfrak{t}$ and $1_G\in G$ in the direction of $\eta\in\mathfrak{t}^\perp$, the change in $1_G t 1_G^{-1}$ to first order is

$$\rho = (1+\eta)t(1+\xi)(1-\eta) - t = t\xi + \eta t - t\eta,$$

so $t^{-1}\rho = \xi + (t^{-1}\eta t - \eta) \in \mathfrak{t}\oplus\mathfrak{t}^\perp = \mathfrak{g}$, from which taking determinants yields $J(t) = \det\big(\mathrm{id}_{\mathfrak{t}}\oplus(A(t^{-1}) - \mathrm{id}_{\mathfrak{t}^\perp})\big) = \det\big(A(t^{-1}) - \mathrm{id}_{\mathfrak{t}^\perp}\big)$, where A is the adjoint action of T on $\mathfrak{t}^\perp$ (obtained by restricting the adjoint action of G on $\mathfrak{g}$ to T on $\mathfrak{t}^\perp$).

Since $\mathfrak{g}$ is spanned by $i = \left(\begin{smallmatrix} i & 0\\ 0 & -i\end{smallmatrix}\right)$, $j = \left(\begin{smallmatrix} 0 & 1\\ -1 & 0\end{smallmatrix}\right)$, and $k = \left(\begin{smallmatrix} 0 & i\\ i & 0\end{smallmatrix}\right)$ we have that $\mathfrak{t}^\perp$ is spanned by j,k. The adjoint action of T on $\mathfrak{t}^\perp$ is given by conjugation by elements of T. If $t = \left(\begin{smallmatrix} e^{is} & 0\\ 0 & e^{-is}\end{smallmatrix}\right)$,

$$tjt^{-1} = \left(\begin{smallmatrix} 0 & e^{2is}\\ -e^{-2is} & 0\end{smallmatrix}\right) = \cos(2s)j + \sin(2s)k \quad\text{and}\quad tkt^{-1} = \left(\begin{smallmatrix} 0 & ie^{2is}\\ ie^{-2is} & 0\end{smallmatrix}\right) = -\sin(2s)j + \cos(2s)k$$

then $A(t) = \left(\begin{smallmatrix} \cos(2s) & -\sin(2s)\\ \sin(2s) & \cos(2s)\end{smallmatrix}\right)$. It follows that $J(t) = \det\big(A(t^{-1}) - \mathrm{id}_{\mathfrak{t}^\perp}\big)$ is $4\sin^2(s) = \left|e^{is} - e^{-is}\right|^2 = |\Delta(t)|^2$ where $\Delta(t)$ is the Vandermonde determinant $\det\left(\begin{smallmatrix} 1 & 1\\ e^{-is} & e^{is}\end{smallmatrix}\right)$. So for a class function f, the Weyl integration formula is $\int_G f\mathrm{d}g = \frac{1}{2}\int_T f(t)|\Delta(t)|^2\mathrm{d}t$. Notice that the normalizations we chose were all correct; that is, take f to be the constant 1 function and see that we do get equality. This is because the Haar measures we chose for G and T were normalized Haar measures.

Let V,W be unitary irreps of G so that $\langle\chi_V,\chi_W\rangle_{L^2(G)} = \delta_{V,W}$. We turn the integral over the group G to an integral over the torus T using the Weyl integration formula, since characters are class functions. We have

$$\delta_{V,W} = \langle\chi_V,\chi_W\rangle_{L^2(G)} = \int_G \chi_V\overline{\chi_W}\mathrm{d}g = \frac{1}{2}\int_T \chi_V\Delta\overline{\chi_W\Delta}\mathrm{d}t = \frac{1}{2}\langle\chi_V\Delta,\chi_W\Delta\rangle_{L^2(T)}$$

We proceed with the calculation in the case that $V = W$. Change coordinates from the torus T to $\mathrm{U}(1)$ (the Jacobian of this transformation is 1). Then $\chi_V(t) = \chi_V(z) = \sum_{p\in\mathbb{Z}} m_p z^p$ is a finite sum and is palindromic; that is, $m_p = m_{-p}$. Then $\Delta(z) = z - z^{-1}$ so that $\Delta(z)\chi_V(z) = \sum_{p\in\mathbb{Z}} m_p z^{p+1} - m_p z^{p-1} = \sum_p (m_{p-1} - m_{p+1})z^p$. This sum is anti-palindromic; that is, if $a_p = m_{p-1} - m_{p+1}$, then $a_p = -a_{-p}$.

We have

$$
\begin{aligned}
1 &= \frac{1}{2}\int_T |\Delta\chi_V|^2 \mathrm{d}t \\
&= \frac{1}{2}\int_{\mathrm{U}(1)} \Delta\chi_V \overline{\Delta\chi_V}\mathrm{d}z \\
&= \frac{1}{2}\int_{\mathrm{U}(1)} \left(\sum_{p\in\mathbb{Z}} a_p z^p\right)\left(\sum_{q\in\mathbb{Z}} \overline{a_q z^q}\right)\mathrm{d}z \\
&\overset{(4)}{=} \frac{1}{2}\int_{\mathrm{U}(1)} \left(\sum_{p\in\mathbb{Z}} a_p z^p\right)\left(\sum_{q\in\mathbb{Z}} a_q z^{-q}\right)\mathrm{d}z \\
&= \frac{1}{2}\int_{\mathrm{U}(1)} \sum_{p,q\in\mathbb{Z}} a_p a_q z^{p-q}\mathrm{d}z \\
&\overset{(6)}{=} \frac{1}{2}\sum_{p\in\mathbb{Z}} a_p^2 \\
&\overset{(7)}{=} \sum_{p=1}^{\infty} a_p^2,
\end{aligned}
$$

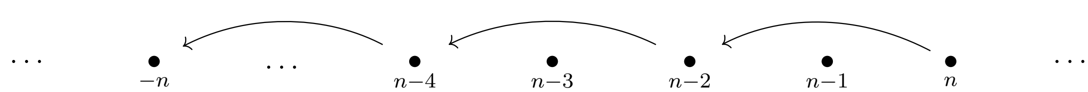
where equality (4) holds since the m_p are integers and hence also the a_p; equalities (6), (7) (haha six-seven hahaha *vomits*; see South Park) hold because $\int_{\mathrm{U}(1)} z^{p-q}\mathrm{d}z = \delta_{pq}$ and because the a_p satisfy $a_p = -a_{-p}$. From this string of equalities conclude that only one of the a_p must be nonzero; say for some n, a_{n+1} is nonzero with $a_{n+1} = \pm 1$ (note also $a_{-n-1} = -a_{n+1} = \mp 1$ as a result). Thus $\chi_V(z) = \Delta(z)\chi_V(z)/\Delta(z) = \sum_{p\in\mathbb{Z}} a_p z^p/(z-z^{-1}) = a_{n+1}(z^{n+1} - z^{-n-1})/(z-z^{-1}) = a_{n+1}\sum_{p=0}^{\infty} z^{n-2p}$. Since the m_p are nonnegative, it follows that a_{n+1} is 1.

We have proved the Weyl character formula: For any nonnegative integer n, there exists an irreducible representation V_n of $G = \mathrm{SU}(2)$ with character

$$\chi_{V_n}(z) = \frac{z^{n+1} - z^{-n-1}}{z - z^{-1}} = z^n + z^{n-2} + \cdots + z^{-n+2} + z^{-n}$$

and that characters of irreps of G occur in this way. Furthermore, the dimension of V_n is $\chi_{V_n}(1) = n+1$. This allows us to deduce that the expression $z^n + 72z^{n-2} + 46z^{n-4} + 46z^{-n+4} + 72z^{-n+2} + z^{-n}$ from before could not have been a character of a representation of G, since even though it is palindromic, the coefficients could not be obtained by adding together the characters of irreps of G.

Typically V_n is drawn as a collection of nodes labeled by integers. On the line of integers, V_n looks like



In the picture the arrows do have meaning, but we will come to them later. There are arrows which do go in the opposite direction as well, as we will see. We think of each node as being a one-dimensional vector space, but which ones and its significance will also come later.

We still do not have a description of each V_n; that will come soon, but at least for now we know their characters. From the Peter-Weyl theorem, we can think of the "dual" of $G = \mathrm{SU}(2)$ as $\mathbb{N}$ with a counting measure whose weight is $1/\dim$, which will end up being $1/n$ for each V_n.

Introduction to Lie algebras

We would like to do more algebra and geometry to study representation theory, so we start by looking at Lie algebras of the Lie groups we've been studying. This will also lead us to understanding what those arrows were in the above picture of V_n; they will end up being elements of the Lie algebra of $G = \mathrm{SU}(2)$ viewed as operators on V_n.

Let G be any Lie group. The group algebra $\mathbb{C}[G]$ acts on G, but there is more that also acts on G, since G is smooth. A Lie algebra is essentially an algebraic structure of vector fields on a manifold. A vector field on a manifold is a first-order differential operator with no constant term; that is, the first-order differential operators which annihilate constant functions. In local coordinates x_i, a vector field looks like $\sum_i f_i \frac{\partial}{\partial x_i}$. A purely algebraic definition of a vector field is as a derivation on the space of smooth functions.

To elaborate, if $\mathcal{O}$ denotes the space of smooth functions on a manifold, we denote by $\mathrm{Der}(\mathcal{O})$ the subspace of endomorphisms of $\mathcal{O}$ which satisfy the Leibniz (product) rule; that is,

$$\mathrm{Der}(\mathcal{O}) = \{\xi \in \mathrm{End}(\mathcal{O}) \mid \xi(fg) = \xi f g + f \xi g\}$$

Alternatively the Leibniz rule can be formulated as a commutation relation in $\mathrm{End}(\mathcal{O})$: we have $[\xi\cdot, f\cdot] = \xi \cdot f \cdot - f \cdot \xi \cdot = (\xi f)\cdot$ where here we embed $\mathcal{O}$ in $\mathrm{End}(\mathcal{O})$ by taking f to pointwise multiplication by f. Notice that derivations are zero on constants; to see this, apply the product rule to $1 = 1 \cdot 1$.

From calculus we know that mixed partial derivatives of smooth enough functions can be taken in any order. The same result should hold for the action of vector fields on functions; that is, for vector fields ξ, η, the endomorphism $[\xi, \eta] = \xi\eta - \eta\xi$ is actually also a vector field itself (it does not act by any second-order differential operators or higher).

A Lie algebra over a field k is a vector space $\mathfrak{g}$ over k equipped with a skew-symmetric bilinear map $[-,-]\colon \bigwedge^2 \mathfrak{g} \to \mathfrak{g}$ satisfying the Jacobi identity

$$[\xi, [\eta, \chi]] = [[\xi, \eta], \chi] + [\eta, [\xi, \chi]]$$

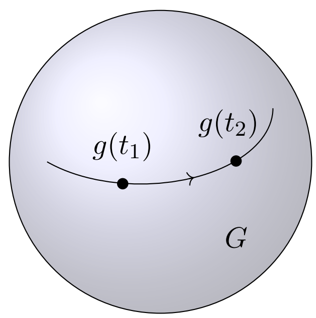
Roughly speaking, $[-,-]$ is a "derivation of itself". We will see later why this condition is natural.

So $\mathrm{Der}(\mathcal{O})$ with Lie bracket $[-,-]$ the commutator of operators is a prototypical Lie algebra. In general we should think of Lie algebras as linearizations of Lie groups, and these are how they occur in nature. There is a functor Lie from the category of Lie groups to Lie algebras taking G to its Lie algebra $\mathfrak{g}$ which we will define in the future. On morphisms, a map of Lie groups $G \xrightarrow{f} H$ is sent to its differential $\mathfrak{g} \xrightarrow{Df} \mathfrak{h}$. This functor is an equivalence of categories when restricted to the subcategory of simply connected Lie groups. This is relevant because SU(2) is a simply connected (compact) Lie group.

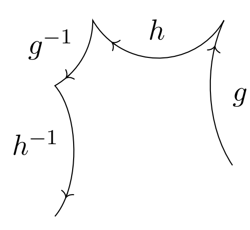
It is reasonable to deal with only the connected Lie groups, but to restrict further to the simply connected Lie groups is not that much of an ask. This is because the universal cover of a connected Lie group is a simply connected Lie group with the same Lie algebra.

We give three definitions of the Lie algebra of a real Lie group G.

1. $\mathrm{Lie}(G) = \mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$. This is to say that the Lie functor is represented by the Lie group $\mathbb{R}$. The elements of $\mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$ are called one-parameter subgroups of G, since the image of a homomorphism of Lie groups $\mathbb{R} \to G$ is a subgroup of G, parameterized by e.g., $t \in \mathbb{R}$. For $g \in \mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$ the picture of a one-parameter subgroup looks like

   where $g(t_1 + t_2) = g(t_1)g(t_2)$. The vector space structure is annoying to write down. It makes sense to define the scalar multiplication by $\lambda g = g(\lambda -)$ but defining $g + h$ is not so easy to guess (and it is annoying to write down). This is one of the downsides of this definition. The bracket $[-,-]$ of maps in $\mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$ is given by $[g,h](t) = g(t)h(t)g^{-1}(t)h^{-1}(t)$. Without supplying a definition for the addition of maps in $\mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$, it will not be possible to verify that this bracket satisfies the Jacobi identity. The commutator looks like

   evaluated at some fixed time t. As t increases, the terminal point should move along the curve defined by the commutator.

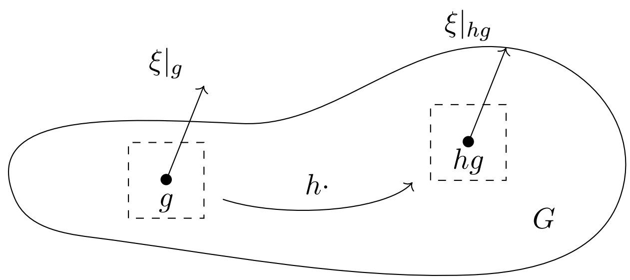
2. Following the discussion about derivations from earlier, one might think to define $\mathrm{Lie}(G)$ as $\mathrm{Der}(C^\infty(G))$, which can be interpreted as the space of vector fields on G. But this space is infinite-dimensional, which will be intractable for our uses. So instead define $\mathrm{Lie}(G)$ by the space of left-invariant vector fields; that

is, by $\mathrm{Der}(C^\infty(G))^G$ where G acts on a derivation/vector field ξ by left translation. The action is given by $g\xi = \xi(g^{-1}-)$, and the left-invariant vector fields are those that are fixed under the left action. By passing to the left-invariant vector fields, we do obtain a finite-dimensional vector space.

For example, on $G = \mathbb{R}^n$, the left-invariant (and simultaneously right-invariant) derivations are the constant coefficient first order differential operators, which are spanned by the differential operators $\frac{\partial}{\partial x_i}$.

The bracket $[-,-]$ on $\mathrm{Der}(C^\infty(G))$ can be restricted to $\mathrm{Der}(C^\infty(G))^G$; that is, the bracket (commutator) of left-invariant derivations is again left-invariant. This is because the Lie bracket of vector fields is coordinate-independent, so left-invariance of the Lie bracket of two left-invariant vector fields follows automatically. Henceforth, we denote the space of left-invariant derivations/vector fields by $\mathfrak{X}(G)$.

What does the left-invariant condition look like? At a point g we can look at a vector field ξ at g to obtain an element $\xi|_g \in T_g(G)$ in Euclidean space (a tangent vector). That ξ is left invariant means that for any $h \in G$, $D(h\cdot)_g \xi|_g$ agrees with $\xi|_{hg}$ in $T_{hg}(G)$.



3. The last definition is the most concrete: The Lie algebra of G is the tangent space of G at the identity 1_G, $T_{1_G}G$. The tangent space of G at the identity is the space of smooth maps from $\mathbb{R}$ to G where 0 is sent to 1_G up to second order and higher (the space of 1-jets at 1_G).

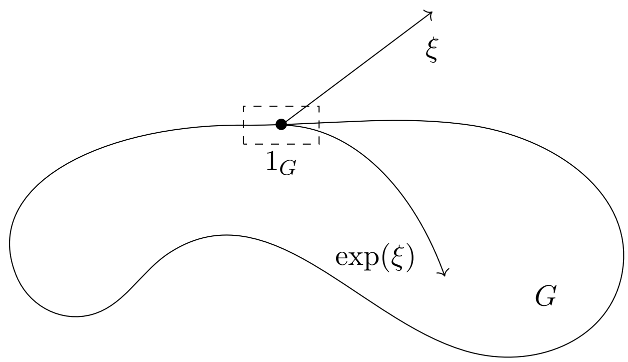
The three definitions above are all equivalent to each other by various maps. We will investigate some of these maps now, and some later.

$$\begin{tikzcd}
 & T_{1_G}(G) \arrow[dl, shift left, "D(\cdot h)\ \forall h"] \arrow[dr, shift left, "\exp"] & \\
\mathfrak{X}(G)^G \arrow[ur, shift left, "(-)|_{1_G}"] \arrow[rr, shift left, "\text{solve ODE}"] & & \mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G) \arrow[ul, shift left, "\frac{\mathrm{d}}{\mathrm{d}t}(-)|_{t=0}"'] \arrow[ll, shift left]
\end{tikzcd}$$

For our purposes, we should not think too hard about Lie algebras of Lie groups, and focus on what they do at a more intuitive level maybe.

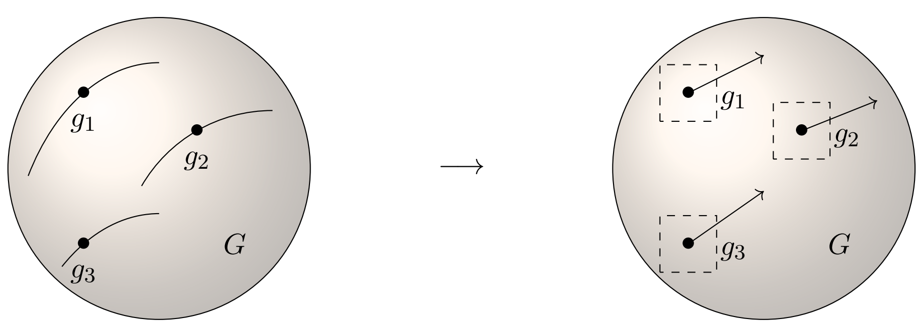
The map $(-)|_{1_G}$ restricts a vector field at the identity of G. The "map" $D(\cdot h)\ \forall h$ means we take a tangent vector at the identity of G and smear it around the entire Lie group by the differential of the right translation map $\cdot h$ for all $h \in G$. Since the action of right translation commutes with the action of left translation on the group G, smearing a tangent vector in $T_{1_G}(G)$ around by right translation will produce a left-invariant vector field.

The map $\frac{\mathrm{d}}{\mathrm{d}t}(-)|_{t=0}$ takes a one-parameter subgroup, which we think of as a curve passing through 1_G, and differentiates it at $t = 0$ to produce a tangent vector at 1_G. The exponential map is a special map which takes a tangent vector v and produces a curve in G passing through 1_G whose derivative at $t = 0$ is v; when G is a matrix group, the exponential map is given by taking a matrix A in $T_{1_G}(G)$ and evaluating $\exp(tA)$ using the Taylor series for exp. The group commutator of curves given by taking exponentials of Lie algebras looks like the earlier picture above with the commutator of one-parameter subgroups, but involves the Lie algebra bracket, by the Baker-Campbell-Hausdorff formula: $\exp(t\xi)\exp(t\eta) = \exp\bigl(\xi + \eta + \frac{1}{2}t^2[\xi, \eta] + \cdots\bigr)$ (see Wikipedia). A picture for the exponential map might look like



The picture above depicts the map $\xi \mapsto \exp(t\xi)$ evaluated at $t = 1$. (See Wikipedia for more details about the exponential map.)

The map labeled "solve ODE" means we solve the differential equation the vector field determines in a neighborhood of 1_G; this is guaranteed by the Picard-Lindelöf existence and uniqueness theorem for a solution to an initial value problem. This looks like starting at 1_G and following the flow of the vector field in a small neighborhood around 1_G (like a paper boat in a river). There are two choices for the unlabeled map at the bottom of the diagram. One choice is to translate a map $t \mapsto \varphi(t)$ around by left multiplying by g for all $g \in G$ to get a family of curves passing through each point of G; then differentiating each of these curves at $t = 0$ to obtain tangent vectors at each $g \in G$ which are assembled into a left-invariant vector field. On the other hand, we can compose the differentiation and right-translation maps to get the same result. The former looks like (at least for just a few points)



Next time we will discuss representations of Lie algebras, specifically of the Lie algebra $\mathfrak{su}(2)$ of the simply connected Lie group $\mathrm{SU}(2)$. A description of $\mathfrak{su}(2)$ is as the space of trace-free, skew-Hermitian complex 2×2 matrices. In what follows, the point is that representations of Lie groups are also representations of the corresponding Lie algebra; that is, the Lie algebra acts infinitesimally on those representations. Since the Lie functor is one-to-one for simply connected Lie groups, we can recover the Lie group representation theory of $\mathrm{SU}(2)$ by studying the Lie algebra representation theory of $\mathfrak{su}(2)$.

Lecture 12 October 02

More on Lie algebras

The Lie bracket of vector fields (derivations) on a Lie group G we defined last time as just the commutator of the operators they define; that is, if ξ, η are vector fields, then their bracket is $[\xi, \eta] = \xi\eta - \eta\xi$, which is also a vector field. Two other ways we can think about the commutator are as follows:

1. We can write $[\xi, \eta]$ as $L_\xi \eta$, where L_ξ is the Lie derivative of ξ. The meaning of this notation: For any $g \in G$, start by solving the ODE that ξ defines in a neighborhood of g to obtain a flow $\Phi_{\xi,g}$ passing through g. If you do this for every $g \in G$, we obtain a self-diffeomorphism of G which we call Φ_ξ; we will suppress notating the g in $\Phi_{\xi,g}$ to refer to the specific flow passing through $g \in G$. We want to "differentiate" η along paths passing through each $g \in G$; that is, for each g we want to compare the tangent vector $\eta|_g = \eta|_{\Phi_\xi(0)} \in T_g G$ with $\eta|_{\Phi_\xi(t)}$ for some small $t > 0$. These vectors belong to different vector spaces, so we will need to have some notion of "parallel transport" to move $\eta|_{\Phi_\xi(t)} \in T_{\Phi_\xi(t)} G$ along the curve Φ_ξ (backwards in t) to a vector in $T_g G$. Then only will it make sense to take the difference of these vectors. picture

   The formula for the Lie derivative $L_\xi \eta$ is
   $$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{d}\Phi_\xi(-t)\eta|_{\Phi_\xi(t)}\Big)|_{t=0}$$
   where $\mathrm{d}\Phi_\xi(-t)$ is meant to be the map which transports tangent vectors in $T_{\Phi_\xi(t)}G$ to $T_g G$ "backwards" along the path defined by Φ_ξ passing through g. We will not think about this notion again, so refer to any text on smooth manifolds to see why the Lie derivative of vector fields coincides with the Lie bracket of vector fields.

2. Let Φ_ξ and Φ_η be the flows that ξ, η define. Then we can take the commutator (of diffeomorphisms) of these flows to get the flow defined by $\Phi_\eta^{-1} \circ \Phi_\xi^{-1} \circ \Phi_\eta \circ \Phi_\xi$. It turns out that the vector field this flow corresponds to is $[\xi, \eta]$; that is, $\Phi_{[\xi,\eta]} = \Phi_\eta^{-1} \circ \Phi_\xi^{-1} \circ \Phi_\eta \circ \Phi_\xi$, and at any $g \in G$ we have
   $$[\xi, \eta]|_g = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big((\Phi_\eta^{-1} \circ \Phi_\xi^{-1} \circ \Phi_\eta \circ \Phi_\xi)(t)\Big)|_{t=0}$$
   where indeed the first derivative of the composite flow vanishes. Again, see other sources for the proofs. picture

Recall from last time that for any Lie group G we defined the Lie algebra of G, denoted $\mathrm{Lie}(G) = \mathfrak{g}$, by three equivalent objects: The tangent space of G at the identity, the space of left-invariant vector fields on G, and as $\mathrm{Hom}_{\mathbf{LieGrp}}(\mathbb{R}, G)$. We also stated last time that the Lie functor sending a Lie group to its Lie algebra defines an equivalence of categories between the category of simply connected Lie groups and the category of Lie algebras. This tells us that not every Lie algebra occurs as the Lie algebra of a Lie group. So what is another way to think of Lie algebras if not from Lie groups?

Lie algebras can also come from associative algebras over fields. If A is an associative algebra over a field k, then define the Lie bracket on A to be $[a, b] = ab - ba$. This defines a Lie algebra structure on A. For example, on a manifold M we can consider the space of smooth functions $C^\infty(M)$ on M and consider the associative algebra $\mathrm{End}(C^\infty(M))$, or equivalently the space of jets of functions on M. Equip this algebra with its natural Lie algebra structure. This is a huge, infinite-dimensional algebra, but the derivations (or for the space of jets, the vector fields) form a Lie subalgebra of $\mathrm{End}(C^\infty(M))$ (or jets) which we usually think about (these are the "first-order" elements). In a similar way, A can be thought of as the "first-order" elements in $\mathrm{End}(A)$, where $\mathrm{End}(A)$ is also an associative algebra with Lie bracket the usual commutator of operators, and A embeds as a Lie subalgebra of $\mathrm{End}(A)$ by $a \mapsto a \cdot -$.

Another nice example: Let V be a n-dimensional vector space over $\mathbb{R}$. Then $\mathrm{End}(V)$ with the usual commutator of operators is the Lie algebra of the associative algebra $\mathrm{Aut}(V) = \mathrm{GL}(V)$. The notation we use for this example is $\mathrm{End}(V) = \mathfrak{gl}(V) = \mathrm{Lie}(\mathrm{GL}(V))$, or if we fix a basis of V, then we can also write $\mathrm{Mat}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R}) = \mathrm{Lie}(\mathrm{GL}_n(\mathbb{R}))$ (where $\mathrm{Mat}_n(\mathbb{R})$ denotes $n \times n$ matrices over $\mathbb{R}$). Since $\mathrm{GL}(V)$ is a Lie group, we can ask about the exponential map $\exp\colon \mathfrak{gl}(V) \to \mathrm{GL}(V)$, which in this case is given by the usual Taylor series for exp (and calculations from analysis make sense of such a series to converge for matrices). For $x \in T_{1_{\mathrm{GL}_n(\mathbb{R})}} \mathrm{GL}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R}) = \mathrm{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ and $t \in \mathbb{R}$, $\exp(tx) = 1 + tx + \frac{t^2x^2}{2} + \cdots$. If we don't want to think about convergence, we can do things the algebraic geometry way and only think about the truncated Taylor series for exp to arbitrary order. The commutator in $\mathfrak{gl}_n(\mathbb{R})$ is the usual matrix commutator.

As usual with objects in mathematics, we should study maps into and out of them. In this class we are concerned with representations, so we define what a representation of a Lie algebra is, starting with some motivation. Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. Then a finite-dimensional representation V of G is given by a map $G \xrightarrow{\rho} \mathrm{Aut}(V)$, where $\mathrm{GL}(V)$ is given some smooth structure. We differentiate ρ to obtain a map $\mathfrak{g} \xrightarrow{\mathrm{d}\rho} \mathfrak{gl}(V)$ of Lie algebras. So in general define a representation of a Lie algebra $\mathfrak{g}$ to be a vector space V and a Lie algebra map $\mathfrak{g} \to \mathfrak{gl}(V)$ which defines how $\mathfrak{g}$ acts on V. Note that the action is by endomorphisms and not automorphisms.

We briefly mention a few well-known Lie algebras of Lie groups: The Lie group $\mathrm{SL}_n(\mathbb{R})$ is the kernel of the determinant map $\mathrm{GL}_n(\mathbb{R}) \xrightarrow{\det} \mathrm{GL}_1(\mathbb{R}) \cong \mathbb{R}$. It is true that the Lie functor is left-exact (one way to see this is to recall that the Lie functor is represented by $\mathbb{R}$), so we can differentiate the determinant map and take the kernel of the resulting map to obtain a description of $\mathfrak{sl}_n(\mathbb{R})$. One can show that the derivative of the determinant map is the trace map (which is a map of Lie algebras $\mathfrak{gl}_n(\mathbb{R}) \to \mathbb{R}$), so the kernel of the trace map is the collection of trace-free matrices. It is also true that for any matrix A, $\frac{\mathrm{d}}{\mathrm{d}t}(\det \exp(tA))|_{t=0} = \operatorname{tr} A$. Therefore $\mathfrak{sl}_n(\mathbb{R})$ is the collection of trace-free real-valued matrices with commutator the usual commutator of matrices. The calculation above can also be repeated for $\mathrm{SL}_n(\mathbb{C})$ to find that $\mathfrak{sl}_n(\mathbb{C})$ is the collection of trace-free complex-valued matrices. This is because $\mathbb{C} \cong \mathbb{R}^2$, so we can do the appropriate calculations in real coordinates.

The Lie group $\mathrm{U}(n) \subset \mathrm{GL}_n(\mathbb{C})$ is characterized by the collection of complex-valued matrices K such that $K\overline{K}^\top = 1_{\mathrm{U}(n)}$. Differentiate the characteristic equation $K\overline{K}^\top = 1_{\mathrm{U}(n)}$ to find that $\mathrm{u}(n) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid A + \overline{A}^\top\}$, the skew-Hermitian complex-valued matrices. It follows that for $\mathrm{SU}(n)$, its Lie algebra $\mathfrak{su}(n)$ is the collection of trace-free and skew-Hermitian matrices.

Two Lie groups can have the same Lie algebra. The Lie groups $\mathrm{O}_n(\mathbb{R})$ and $\mathrm{SO}_n(\mathbb{R})$ have Lie algebras the collection of skew-symmetric real-valued matrices. This is because the condition of being trace-free is subsumed in the skew-symmetry. Some Lie algebras are already familiar. The Lie algebra of $\mathrm{SO}_3(\mathbb{R})$ can be identified with $\mathbb{R}^3$ equipped with the cross product as the Lie bracket. It is also no coincidence that the same identification can be made for the Lie algebra of $\mathrm{SU}(2)$, given by the $\mathbb{R}$-span of $\{i = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right), j = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), k = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)\}$ with Lie bracket the ordinary product of matrices. It follows that if V is a finite-dimensional Lie algebra representation of $\mathfrak{su}(2)$; that is, there is a map of Lie algebras $\mathfrak{su}(2) \xrightarrow{\varphi} \mathrm{End}(V)$ with $[\varphi(i), \varphi(j)] = \varphi(k)$, $[\varphi(j), \varphi(k)] = \varphi(i)$, and other relations coming from the multiplication law in $\mathfrak{su}(2)$.

For G simply connected, because we can recover Lie groups from their Lie algebras, it is further true that the finite-dimensional complex Lie algebra representations of $\mathrm{Lie}(G)$ are in bijection with the finite-dimensional complex Lie group representations of G. This is because the elements of each side are parameterized by maps $\mathrm{Lie}(G) \to \mathfrak{gl}_n(\mathbb{C})$ and $G \to \mathrm{GL}_n(\mathbb{C})$ respectively, and we can pass between these collections by exponentiation and differentiation. A more general theorem is that if G, H are Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$ respectively, a Lie algebra homomorphism $\mathfrak{g} \xrightarrow{\phi} \mathfrak{h}$ exponentiates uniquely to a Lie group homomorphism $G \xrightarrow{\Phi} H$ with $\Phi(e^x) = e^{\phi(x)}$ for all $x \in \mathfrak{g}$ if G is simply connected.

Weyl's unitary trick

The group $G = \mathrm{SL}_2(\mathbb{C})$ is simply connected, since topologically G is isomorphic to $S^3 \times \mathbb{R}$ (the S^3 factor is $\mathrm{SU}(2)$ and the $\mathbb{R}$ is there to scale the determinant). Let $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{C}))$ denote the category of finite-dimensional holomorphic complex Lie group representations of $\mathrm{SL}_2(\mathbb{C})$, and $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C}))$ is the category of finite-dimensional complex Lie algebra representations of $\mathfrak{sl}_2(\mathbb{C})$. We want to show that $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{C}))$ is equivalent to $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C}))$, but the adjective "holomorphic" is questionable. Exponentiating a representation of $\mathfrak{sl}_2(\mathbb{C})$ to a representation of $\mathrm{SL}_2(\mathbb{C})$ produces a holomorphic representation (since a map of complex Lie algebras is complex-linear), so this gives the desired equivalence.

A similar result is true for $G = \mathrm{SU}(2)$, where $\mathbf{Rep}(\mathrm{SU}(2))$ is equivalent to $\mathbf{Rep}(\mathfrak{su}(2))$; here we only require the representations in $\mathbf{Rep}(\mathrm{SU}(2))$ to be representations of Lie groups. Note that since $\mathrm{SU}(2)$ is compact, the representations in $\mathbf{Rep}(\mathrm{SU}(2))$ are unitarizable. For $\mathrm{SU}(2)$, the equivalence is more concrete. Every element of $\mathrm{SU}(2)$ belongs to some torus, but observe that every torus in $\mathrm{SU}(2)$ is a one-parameter subgroup; that is, every torus occurs as $\{\exp(tA) \mid t \in \mathbb{R}\}$ for $A \in \mathfrak{su}(2)$ (this is the usual polar decomposition for $\mathrm{SU}(2)$). In other words, the exponential map $\mathfrak{su}(2) \xrightarrow{\exp} \mathrm{SU}(2)$ is surjective. (In fact, the exponential map $\mathrm{Lie}(G) \xrightarrow{\exp} G$ is surjective for G a connected compact Lie group.) So any Lie algebra representation $\mathfrak{su}(2) \xrightarrow{\phi} \mathrm{End}(V)$ exponentiates uniquely to a Lie group representation $\mathrm{SU}(2) \xrightarrow{\Phi} \mathrm{GL}(V)$ and differentiation gives the correspondence in the other direction.

The "unitary trick" is the following theorem: that the collections of

1. (smooth) finite-dimensional representations of $\mathrm{SU}(2)$,
2. finite-dimensional representations of $\mathfrak{su}(2)$,
3. holomorphic finite-dimensional representations of $\mathrm{SL}_2(\mathbb{C})$,
4. representations of $\mathfrak{sl}_2(\mathbb{C})$ (specifically these are given by maps of complex Lie algebras),
5. (smooth) finite-dimensional representations of $\mathrm{SL}_2(\mathbb{R})$, and
6. finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{R})$

are equivalent to each other; that is, there are bijections between these collections which preserve invariant subspaces and isomorphisms of representations. We will see in the future that the finite-dimensional representations of $\mathrm{SU}(2)$ and $\mathrm{SL}_2(\mathbb{R})$ were already smooth (meaning we don't need that adjective above in 1. and 5.). This explains why, throughout the notes, we have not been so careful with this adjective.

To see why this theorem is true, we start with a $\mathfrak{g}$ a Lie algebra over $\mathbb{R}$. An n-dimensional complex representation of $\mathfrak{g}$ is given by an $\mathbb{R}$-linear map $\mathfrak{g} \xrightarrow{\phi} \mathrm{Mat}_n(\mathbb{C})$. Since the target $\mathrm{Mat}_n(\mathbb{C})$ is a $\mathbb{C}$-vector space, we may extend ϕ uniquely to a $\mathbb{C}$-linear map $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} \to \mathrm{Mat}_n(\mathbb{C})$ by the universal property of the tensor product.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\phi"] \arrow[d] & \mathrm{Mat}_n(\mathbb{C}) \\ \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} \arrow[ur, dashed] & \end{tikzcd}$$

The object $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ is denoted $\mathfrak{g}_{\mathbb{C}}$, and is called the complexification of the real Lie algebra $\mathfrak{g}$. The complex vector space $\mathfrak{g}_{\mathbb{C}}$ is given the structure of a complex Lie algebra by extending the bracket on $\mathfrak{g}$ to a bracket on $\mathfrak{g}_{\mathbb{C}}$ by $\mathbb{C}$-linearity, and $\mathfrak{g}$ embeds as a real Lie subalgebra of $\mathfrak{g}_{\mathbb{C}}$. If $\mathfrak{g}$ is the Lie algebra of a (real) Lie group G, then it is the case that $\mathfrak{g}_{\mathbb{C}}$ is the Lie algebra of the complex Lie group (i.e. viewed as a complex manifold) $G_{\mathbb{C}}$, which is the complexification of G. We will not think about $G_{\mathbb{C}}$ in what follows. In the above observation, we merely extended the Lie algebra homomorphism ϕ to a map of complex vector spaces, but because the bracket on $\mathfrak{g}_{\mathbb{C}}$ is obtained by extending the bracket on $\mathfrak{g}$ $\mathbb{C}$-linearly, the extended map $\mathfrak{g}_{\mathbb{C}} \to \mathrm{Mat}_n(\mathbb{C})$ is a map of complex (more than just real!) Lie algebras. This shows that $\mathbf{Rep}(\mathfrak{g})$ is equivalent to $\mathbf{Rep}(\mathfrak{g}_{\mathbb{C}})$ (where again, the maps defining representations of the complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$ are complex Lie algebra maps) since the representations in these categories are complex vector spaces.

Recall that our favorite $\mathbb{R}$-basis for $\mathfrak{su}(2)$ is given by $\{i = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right), j = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), k = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)\}$, forming a real vector subspace of the complex vector space $\mathrm{Mat}_2(\mathbb{C})$. By complexifying, it follows that $\mathfrak{su}(2) \otimes_{\mathbb{R}} \mathbb{C}$ is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$, the collection of trace-free complex-valued matrices (as complex Lie algebras). So $\mathbf{Rep}(\mathfrak{su}(2))$ is equivalent to $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C}))$. A small calculation shows that $\mathfrak{sl}_2(\mathbb{R})$ is the collection of trace-free real-valued matrices, so its complexification is $\mathfrak{sl}_2(\mathbb{C})$. So $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{R}))$ is equivalent to $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C}))$. We obtain the following diagram:

$$\begin{tikzcd}
\mathbf{Rep}(\mathfrak{su}(2)) \arrow[r, leftrightarrow, "\cong"] \arrow[d, leftrightarrow, "\cong"'] & \mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C})) \arrow[r, leftrightarrow, "\cong"] \arrow[d, leftrightarrow, "\cong"'] & \mathbf{Rep}(\mathfrak{sl}_2(\mathbb{R})) \\
\mathbf{Rep}(\mathrm{SU}(2)) & \mathbf{Rep}(\mathrm{SL}_2(\mathbb{C})) \arrow[l, "\text{restrict}"] \arrow[r, "\text{restrict}"'] & \mathbf{Rep}(\mathrm{SL}_2(\mathbb{R})) \arrow[u, "\text{differentiate}"']
\end{tikzcd}$$

The vertical arrows in the above diagram form equivalences since $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SU}(2)$ are simply connected, and exponentiating a representation of $\mathfrak{sl}_2(\mathbb{C})$ to a representation of $\mathrm{SL}_2(\mathbb{C})$ produces a holomorphic representation. This establishes most of the equivalences we need, but we are missing the equivalence between $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{R}))$ and the other five nodes above. It turns out that an equivalence between it and $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{C}))$ can be verified directly as we will see much later on, but despite $\mathrm{SL}_2(\mathbb{R})$ not being simply connected, we can still see that $\mathbf{Rep}(\mathfrak{sl}_2(\mathbb{R}))$ and $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{R}))$ are equivalent by taking the counterclockwise loop on the right starting and ending at $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{C}))$. Starting with a representation of $\mathrm{SL}_2(\mathbb{C})$ and restricting, differentiating, complexifying, and exponentiating again will yield the same representation we started with. This is because $\mathrm{SL}_2(\mathbb{R})$ is connected, so a differential of a Lie group homomorphism out of this group determines the homomorphism.

The name "unitary trick" comes from the fact that finite-dimensional representations of $\mathrm{SU}(2)$ are unitarizable. Since we saw that these representations were semisimple, it follows that the representations in the other five collections above are also. (But this is not saying that all of the finite-dimensional representations of $\mathrm{SL}_2(\mathbb{R})$ are unitarizable! In fact, unitary finite-dimensional representations of $\mathrm{SL}_2(\mathbb{R})$ are trivial.) (See Chapter 2, Section 1 of Knapp's Representation Theory of Semisimple Groups for the unitary trick and this fun fact)

So as promised, we will start with the finite-dimensional representation theory of $\mathrm{SU}(2)$, and follow the arrows around to obtain the finite-dimensional (holomorphic) representation theory of $\mathrm{SL}_2(\mathbb{C})$. For us, $\mathrm{SU}(2)$ is the simplest non-Abelian group we care about, but from the point of view of harmonic analysis, the Peter-Weyl theorem justifies our choice in starting with the representation theory of $\mathrm{SU}(2)$, since it is a compact Lie group. There the finite-dimensional representation theory is completely reducible and we found all the irreps already.

A tiny preview of $\mathfrak{sl}_2(\mathbb{C})$

We can obtain a basis of $\mathfrak{sl}_2(\mathbb{C})$ by complexifying the basis $\{i = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right), j = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), k = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)\}$ of $\mathfrak{su}(2)$, but this is not a "democratic" basis for $\mathfrak{sl}_2(\mathbb{C})$ for reasons we will see in future lectures. A more democratic basis for $\mathfrak{sl}_2(\mathbb{C})$ is given by

$$\{e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), f = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\}$$

These elements satisfy the relations

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f$$

Let V be a finite-dimensional representation of $\mathrm{SU}(2)$. By differentiating, obtain a representation of $\mathfrak{su}(2)$, and complexify to obtain a representation of $\mathfrak{sl}_2(\mathbb{C})$. A representation V of $\mathfrak{sl}_2(\mathbb{C})$ amounts to specifying elements

E, H, F in $\mathrm{End}_{\mathbb{C}}(V)$ satisfying the same commutation relations as the ones above for e, h, f. Note that even though we started with a unitarizable representation V of $\mathrm{SU}(2)$, the representation given by the action of $\mathfrak{sl}_2(\mathbb{C})$ on V need not be unitarizable (after all, E, H, F merely belong to $\mathrm{End}_{\mathbb{C}}(V)$).

Returning to $\mathrm{SL}_2(\mathbb{C})$, there is a distinguished subgroup we will look at. Denote by H the subgroup consisting of the diagonal matrices in $\mathrm{SL}_2(\mathbb{C})$; that is, $H = \{\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right) \mid a \in \mathbb{C}^\times\}$. From this description see that $H \cong \mathbb{C}^\times$. By restricting to the copy of $\mathrm{U}(1)$ in $\mathbb{C}^\times$, we can use the representation theory we obtained earlier. In particular, see that $\mathrm{Lie}(H)$ as a Lie subalgebra of $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$ is $\mathbb{C}h$ (differentiate a path in $\mathrm{SL}_2(\mathbb{C})$ taking values in H passing through the identity matrix), which is also the complexification of the real Lie algebra $i\mathbb{R}h$ of the copy of $\mathrm{U}(1)$ in H. From the $\mathrm{U}(1)$-representation theory from earlier, it follows that h acts on a representation semi-simply (meaning the representation restricted to $i\mathbb{R}h$, and hence also $\mathbb{C}h$, is completely reducible).

Lecture 13 October 07

Short peek back at the Peter-Weyl theorem

The approach to start with finite-dimensional representations of SU(2) and follow the arrows in the diagram

$$\begin{tikzcd} \mathbf{Rep}(\mathfrak{su}(2)) \arrow[r, "\cong"] & \mathbf{Rep}(\mathfrak{sl}_2(\mathbb{C})) \arrow[d, "\cong"] \\ \mathbf{Rep}(\mathrm{SU}(2)) \arrow[u, "\cong"] & \mathbf{Rep}(\mathrm{SL}_2(\mathbb{C})) \end{tikzcd}$$

to obtain finite-dimensional representations of $\mathrm{SL}_2(\mathbb{C})$ is comparable to the following "Abelian case" of passing from finite-dimensional representations of U(1) to finite-dimensional representations of $\mathbb{C}^\times$ following arrows in a similar diagram

$$\begin{tikzcd} \mathbf{Rep}(i\mathbb{R}) \arrow[r, "\cong"] & \mathbf{Rep}(\mathbb{C}) \arrow[d, "\cong"] \\ \mathbf{Rep}(\mathrm{U}(1)) \arrow[u, "\cong"] & \mathbf{Rep}(\mathbb{C}^\times) \end{tikzcd}$$

Here the Lie algebras $i\mathbb{R}$ and $\mathbb{C}$ have the trivial (zero) Lie bracket, and note also that $\mathbb{C}^\times \cong \mathrm{U}(1) \times \mathbb{R}_{>0}$ (polar decomposition). The irreps of U(1) are indexed by $n \in \mathbb{Z}$, and the irreps of $\mathbb{R}_{>0} \cong \mathbb{R}$ are indexed by $t \in \mathbb{R}$. In order to match the irreps, we should restrict the kinds of representations appearing in $\mathbf{Rep}(\mathbb{C}^\times)$; that is, $\mathbf{Rep}(\mathbb{C}^\times)$ is the category of finite-dimensional algebraic or holomorphic representations of $\mathbb{C}^\times$. Indeed, a character of $\mathbb{C}^\times$ is a group homomorphism $\mathbb{C}^\times \to \mathbb{C}^\times$, which can either take the forms $z \mapsto z^n$ for $n \in \mathbb{Z}$ or $z \mapsto \overline{z}^n$ for $n \in \mathbb{Z}$. Only the first collection of maps are holomorphic (or algebraic), and these are exactly the same characters of U(1).

Let G be a compact Lie group. In the discussion of the Peter-Weyl theorem, we mentioned that the functions obtained from applying the matrix elements map to $\bigoplus_{V \text{ irrep}} V \oplus V^*$, which we denoted by $C(G)^{\mathrm{fin}}$, coincide with the algebraic functions on G, denoted $\mathbb{C}[G]$. We say that a representation V of G is algebraic if $V = V^{\mathrm{fin}}$. This is equivalent to V decomposing as a direct sum of its irreps (and not as a completion, e.g. in the case of $V = L^2(G)$, V is not algebraic unless G is finite probably). Another equivalent description of an algebraic representation is that the matrix elements of V are algebraic functions. Take $G = \mathrm{U}(1)$ for example; the algebraic functions on G are given by polynomials in $\mathbb{C}[z, z^{-1}]$. The same is true for $G = \mathbb{C}^\times$; that is, $\mathbb{C}[\mathbb{C}^\times] \cong \mathbb{C}[z, z^{-1}]$. The most basic description of an algebraic representation V of G is that the map $G \to \mathrm{GL}(V)$ is an algebraic map.

A fact we will not prove: that an algebraic representation of $\mathbb{C}^\times$ or $\mathrm{SL}_2(\mathbb{C})$ is holomorphic (and vice versa); in other words, that the matrix elements are algebraic functions implies they are holomorphic and vice versa. (This is true of semisimple complex linear algebraic groups according to this MathOverflow post.)

Consider the algebraic (holomorphic) representation $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})] = C(\mathrm{SL}_2(\mathbb{C}))^{\mathrm{fin}}$ of $\mathrm{SL}_2(\mathbb{C})$, which by restricting the functions to SU(2), we obtain an isomorphic ring of functions $\mathbb{C}[\mathrm{SU}(2)] = C(\mathrm{SU}(2))^{\mathrm{fin}}$, which SU(2) acts on naturally. So in view of the diagram of part of the unitary trick above, this is following the not-drawn restriction map from $\mathbf{Rep}(\mathrm{SL}_2(\mathbb{C}))$ to $\mathbf{Rep}(\mathrm{SU}(2))$. A similar thing can be done for $\mathbb{C}[\mathbb{C}^\times]$; restriction of this representation of $\mathbb{C}^\times$ yields the representation $\mathbb{C}[\mathrm{U}(1)]$ of $\mathrm{U}(1)$. That function restriction in both examples is an isomorphism of rings requires proof (e.g. by analysis). As a fun fact, it is possible to define the complexification of a compact Lie group G as $\mathrm{Spec}(\mathbb{C}[G])$; this complexification is a complex algebraic group which is denoted $G_{\mathbb{C}}$ (source??).

Start of highest weight theory of $\mathfrak{sl}_2(\mathbb{C})$

Recall that $\mathfrak{sl}_2(\mathbb{C})$ is the $\mathbb{C}$-span of $\{e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), f = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\}$ satisfying

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f$$

and that $\mathbb{C}h$ is the Lie algebra of the subgroup $H = \{\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right) \mid a \in \mathbb{C}^\times\} \cong \mathbb{C}^\times$ of $\mathrm{SL}_2(\mathbb{C})$, viewed as a Lie subalgebra of $\mathfrak{sl}_2(\mathbb{C})$.

Let V be a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$, and suppose that $v \in V$ is an eigenvector with eigenvalue $\lambda \in \mathbb{C}$ of the action of h on V; that is, $hv = \lambda v$. Then $hev = ehv + [h, e]v = e(\lambda v) + 2ev = (\lambda + 2)ev$. So if $ev \neq 0$, ev is an eigenvector of h with eigenvalue $\lambda + 2$. A similar calculation shows that fv (when it's nonzero) is an eigenvector of h with eigenvalue $\lambda - 2$. The action of h on V need not be diagonalizable, but since V is finite-dimensional, we will obtain finitely many generalized eigenvectors of h (by finding the Jordan normal form for h). However, there will be at least one eigenvector v of h. For an eigenvector v of h, $e^n v = e \cdots ev = 0$ for some n and $f^m v = f \cdots fv = 0$ (where the smallest values for n, m satisfying these equations of course depend on v). Indeed, h can only have finitely many eigenvalues.

By repeated application of e on some eigenvector of h, there exists an eigenvector v of h with $hv = \lambda v$ for which $ev = 0$, which we call a highest weight vector of V with weight λ. There can be more than one highest weight vectors in V; the adjective "highest" refers to highest weight vectors being annihilated by e. Let $v = v_\lambda$ be a highest weight vector of V of weight λ. Then repeated application of f to v produces a finite collection of vectors which are usually drawn in the following way:

$$\begin{array}{ccccccccc} f^m v_\lambda & \xleftarrow{f} & \cdots & \xleftarrow{f} & f^2 v_\lambda & \xleftarrow{f} & f v_\lambda & \xleftarrow{f} & v_\lambda \\ \bullet & & & & \bullet & & \bullet & & \bullet \\ \lambda-2m & & & & \lambda-4 & & \lambda-2 & & \lambda \end{array}$$

Each bullet is labeled above by the vector obtained by repeated application of f on v_λ and the corresponding weight (the h-eigenvalue) is labeled below. For example, $hfv_\lambda = (\lambda - 2)fv_\lambda$. Implicit in the diagram is that $f^{m+1}v_\lambda = 0$. The vector $f^m v_\lambda$ is called a lowest weight vector. Observe further that the subspace of V spanned by $\{f^m v_\lambda, \ldots, v_\lambda\}$ is preserved by the action of h and f. In particular each $f^i v_\lambda$ spans a one-dimensional h-eigenspace of V, and f moves down eigenspaces, lowering the weight. A natural question to ask is if this subspace is preserved by the action of e.

The calculation $efv_\lambda = fev_\lambda + [e, f]v_\lambda = 0 + hv_\lambda = \lambda v_\lambda$ provides some hope. Indeed, the related calculation $ef^i v_\lambda = i(\lambda - i + 1)f^{i-1}v_\lambda$ shows that the action of e on each h-eigenspace moves vectors up weight spaces, but also multiplies by a scalar. So the action of e preserves the subspace spanned by the $f^i v_\lambda$; it follows that this subspace is an irrep of $\mathfrak{sl}_2(\mathbb{C})$

As a corollary, if V is a finite-dimensional irrep of $\mathfrak{sl}_2(\mathbb{C})$, then V is spanned by $\{f^m v_\lambda, \dots, v_\lambda\}$ for some highest weight vector v_λ of weight λ in V. The weight λ in this case is the greatest eigenvalue of the action of h on V. The integer m as before depends on v_λ, and is the dimension of V. To find the irreps of $\mathfrak{sl}_2(\mathbb{C})$, we should first calculate the possible highest weights λ and the possible values of m.

One approach which gets close is to restrict an irrep V of $\mathfrak{sl}_2(\mathbb{C})$ to a representation of $\mathrm{Lie}(H) = \mathbb{C}h$. in this case V would decompose into a direct sum of each of each h-eigenspace, but more importantly we can exponentiate the action of h on V to obtain a representation of $H \cong \mathrm{U}(1)$ on V, where the decomposition of V into each of the h-eigenspaces is preserved as we pass to a representation of $\mathrm{U}(1)$. The action of h on $f^i v_\lambda$ is by multiplication by $\lambda - 2i$, so by exponentiating this action, we obtain the action of $\left(\begin{smallmatrix}\exp(it) & 0 \\ 0 & \exp(-it)\end{smallmatrix}\right) \in H \cong \mathrm{U}(1)$ on $f^i v_\lambda$ by the formula

$$\left(\begin{smallmatrix}\exp(it) & 0 \\ 0 & \exp(-it)\end{smallmatrix}\right) \cdot f^i v_\lambda = \exp(it(\lambda - 2i)) f^i v_\lambda$$

This formula comes from following the maps in the diagram

$$\begin{tikzcd} H \cong \mathrm{U}(1) \arrow[d, "\log"'] & \mathrm{GL}(\mathbb{C} f^i v_\lambda) \\ \mathbb{C}h \arrow[r] & \mathrm{End}(\mathbb{C} f^i v_\lambda) \arrow[u, "\exp"'] \end{tikzcd}$$

In order for the formula above to define an action, $\lambda - 2i$ must be an integer, so λ must be an integer. So only integer weights are permitted for highest weights appearing in irreps of $\mathfrak{sl}_2(\mathbb{C})$.

We can further recall the representation theory of $\mathrm{SU}(2)$ to remember that characters of representations of $\mathrm{SU}(2)$ appear as palindromic polynomials in z, z^{-1}. By passing from a representation of $\mathfrak{sl}_2(\mathbb{C})$ to a representation of $\mathrm{SU}(2)$, it follows that an irrep V with highest weight vector of weight λ should have a lowest weight vector of weight $-\lambda$. So λ can be assumed to be a nonnegative integer. It remains to find the permissible values of $m = \lambda$ (since the action of f reduces the weight by two), which amount to finding which nonnegative integers λ occur for irreps.

The other approach only uses Lie algebras to determine that λ is a nonnegative integer. Let V be an irrep of $\mathfrak{sl}_2(\mathbb{C})$ and let $f^m v_\lambda$ be the lowest weight vector for V. Since it is a lowest weight vector, $f^{m+1} v_\lambda = 0$, hence also $ef^{m+1} v_\lambda = 0$. On the other hand, $ef^{m+1} v_\lambda$ is equal to $(m+1)(\lambda - m) f^m v_\lambda$. Certainly $m+1$ is not zero since m is a nonnegative integer. So $\lambda = m$ as we discovered earlier. We can find examples for small m, but we still don't know which nonnegative integers occur for irreps.

For $m = 0$, the zero vector space does the job. For $m = 1$, take the trivial representation. For $m = 2$, consider $\mathbb{C}^2$ as a representation of $\mathfrak{sl}_2(\mathbb{C})$ where the action is by usual matrix multiplication. The weight spaces in this case are $\mathbb{C}$ with weight -1 and 1. Indeed, the eigenvectors for h in this case are e_2 and e_1 with weights -1 and 1 respectively. For $m = 3$, let $\mathfrak{sl}_2(\mathbb{C})$ act on itself by the Lie bracket; that is, $a \cdot b = [a, b]$. This is called the adjoint representation of $\mathfrak{sl}_2(\mathbb{C})$. In this case the $\mathbb{C}$-spans of f, h, and e form the weight spaces with weights -2, 0, and 2 respectively, from the relations $[h, e] = 2e$ and $[h, f] = -2f$ (h commutes with itself of course). For the other nonnegative integers, we should try a more robust approach.

From our previous lessons with different groups, it may be helpful to search for other representations inside a suitable version of $\mathrm{Fun}(G)$ via the matrix elements map. After all, at least for compact groups, the Peter-Weyl theorem says that for compact groups their representations live inside $L^2(G)$ with multiplicity.

Let V be a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ (and remember by exponentiation V may also be regarded as a representation of $\mathrm{SL}_2(\mathbb{C})$) and consider its dual representation V^* (which is of course isomorphic to V). Pick v_n in V^* to be a highest weight vector with weight n. Exponentiate the action of $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ on V^* to obtain an action of the subgroup $N = \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\} \cong \mathbb{C}$ (note N does not intersect nontrivially with $\mathrm{SU}(2) \subset \mathrm{SL}_2(\mathbb{C})$, and the Lie algebra of N is $\mathbb{C}e$). Since $ev_n = 0$, the action of N on v_n is given by $\left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)v_n = \exp(te)v_n = \sum_{i=0}^{\infty}(te)^i/i!v_n = \sum_{i=0}^{\infty}(t^i/i!)e^i v_n = v_n$. So v_n is N-invariant; that is, $v_n \in (V^*)^N$.

Now restrict the matrix elements map $V \otimes V^* \xrightarrow{f_{-,-}} \mathbb{C}[\mathrm{SL}_2(\mathbb{C})]$ to $V \xrightarrow{f_{-,v_n}} \mathbb{C}[\mathrm{SL}_2(\mathbb{C})]$. The resulting function f_{-,v_n} is $\mathrm{SL}_2(\mathbb{C}) \times \mathrm{SL}_2(\mathbb{C})$-invariant, but because v_n is N-invariant, the functions obtained from f_{-,v_n} are left N-invariant. Indeed, for any $v \in V$, $\left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right) f_{v,v_n}(-) = \langle v_n, \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)^{-1}(-)v\rangle = \langle \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)v_n, (-)v\rangle = \langle v_n, (-)v\rangle = f_{v,v_n}(-)$ (in general, $gf_{v,w}h = f_{hv,gw}$). We will denote the space of left N-invariant polynomial functions on $\mathrm{SL}_2(\mathbb{C})$ by $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^N$.

In general, if X is a geometric space of some kind and some group H acts on $\mathrm{Fun}(X)$, then we may identify the left H-invariant functions, $\mathrm{Fun}(X)^H$, with the orbit space $H\backslash X$. So in our case, we find that $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^N = \mathbb{C}[N\backslash \mathrm{SL}_2(\mathbb{C})]$. Since N is the stabilizer of the vector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $N\backslash \mathrm{SL}_2(\mathbb{C}) = \mathrm{Stab}\{\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\}\backslash \mathrm{SL}_2(\mathbb{C})$ is the orbit of $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ under the action of $\mathrm{SL}_2(\mathbb{C})$, which ends up being $\mathbb{C}^2 \setminus \{0\}$. So $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^N = \mathbb{C}[N\backslash \mathrm{SL}_2(\mathbb{C})] = \mathbb{C}[\mathbb{C}^2 \setminus \{0\}]$. We now cite Hartogs' extension theorem from the theory of functions of several complex variables, stating that a function of more than one complex variable which is holomorphic on a connected set, that is the complement of a compact set, may be extended holomorphically to the compact set. In our case this means that $\mathbb{C}[\mathbb{C}^2 \setminus \{0\}] = \mathbb{C}[\mathbb{C}^2] = \mathbb{C}[x, y]$, the polynomials in two variables with complex coefficients (a (bi)graded ring). In summary, the matrix coefficients map has the signature $V \xrightarrow{f_{-,v_n}} \mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^N = \mathbb{C}[x, y]$. That we used Hartogs' extension theorem is not something that appears in the study of representations of most other groups. So far we have only used the "highest" part of "highest weight theory".

In the case that V is irreducible, the matrix elements map $V \xrightarrow{f_{-,v_n}} \mathbb{C}[x, y]$ is injective so that V may be identified with some subspace of $C[x, y]$. Now we use the "weight" part of "highest weight theory" to proceed. Since $hv_n = nv_n$, we may exponentiate $\mathbb{C}h$ to find that $\mathbb{C}v_n$ is a representation of $H = \{\left(\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix}\right)\} \cong \mathbb{C}^\times$. But $\mathbb{C}v_n$ is also a representation of $N = \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\} \cong \mathbb{C}$, so by taking their semidirect product inside $\mathrm{SL}_2(\mathbb{C})$, $\mathbb{C}v_n$ is a representation of $B = N \rtimes H = \{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\}$. So the matrix elements map f_{-,v_n} injects V into a further refined collection of functions inside of $\mathbb{C}[x, y] = \mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^N \subset \mathbb{C}[\mathrm{SL}_2(\mathbb{C})]$, which we denote by $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})]^{B,n}$. In the next lecture we will define this space and finally write down the finite-dimensional irreps of $\mathrm{SL}_2(\mathbb{C})$.

Lecture 14 October 09

The finite-dimensional irreps of $\mathrm{SL}_2(\mathbb{C})$

References

The references below were from the course page.

Basic Lie theory

R. Carter, G. Segal and I.MacDonald: Lectures on Lie Groups and Lie Algebras. London Math. Society Student Texts 32. The lectures by Segal are a beautiful overview of the fundamental ideas of Lie groups and algebras, with geometry and examples emphasized.

W. Fulton and J. Harris: Representation Theory. Springer GTM 129. The canonical reference for representations, especially for the (Lie) algebraic point of view and basic algebro–geometric aspects.

D. Milicic, The Weyl Character Formula. Explicit derivation for SU(2) not using anything "external" (Part of a course on representation theory.)

Review articles

G. Mackey: Harmonic Analysis as the Exploitation of Symmetry. Bull. Amer. Math. Soc. 3/1 (1980) 543-699. Reprinted in: The Scope and History of Commutative and Noncommutative Harmonic Analysis. History of Math Vol.5, Amer. Math Soc./London Math Soc. 1992. My favorite (and first) introduction to representation theory, emphasizing its history and origins in probability, number theory and physics. The reprint appears in a volume devoted to Mackey's wonderful representation theory survey articles.

Other Mackey surveys (besides those in the book above): Unitary Group Representations in physics, probability and number theory (Addison-Wesley 1989) – a book delving deeper into the ideas of the above survey article.

W. Schmid: Representations of semi-simple Lie groups. In: Atiyah et al., Representation Theory of Lie Groups. Londom Math Society Lecture Notes 34. Cambridge U. Press 1979. Excellent overview by a master, in a volume full of useful reviews (also Mackey, Bott, Kostant, Kazhdan..)

W. Schmid: Analytic and Geometric Realization of Representations. In: Tirao and Wallach (eds)., New Developments in Lie Theory and their Applications. Lecture notes overviewing the subject in the title, with a strong emphasis on SL_2.

R. Howe, On the role of the Heisenberg group in harmonic analysis.

More advanced, SL_2 specific references

K. Vilonen: Representations of SL_2. Course Notes, Northwestern University.

R. Howe and E.C. Tan: Non-abelian harmonic analysis - applications of $\mathrm{SL}(2,\mathbb{R})$. Springer Universitext. A very nice introduction to representations of $\mathrm{SL}(2,\mathbb{R})$, with interesting applications to classical analysis.

V.S. Varadarajan: An Introduction to Harmonic Analysis on Semisimple Lie Groups. Cambridge Studies in advanced math 16. Nice textbook, with an emphasis on SL_2 and good introductions to various topics.

S. Lang: $\mathrm{SL}(2,\mathbb{R})$. Springer GTM. Fairly analytical introduction, alas not very coherent – read the review by Robert Langlands

Other

D. Ramakrishnan and R. Valenza: Fourier Analysis on Number Fields. Springer GTM 186. Contains an introduction to the Pontrjagin duality theory for locally compact abelian groups, and its applications to number theory (Tate's thesis).

J. Arthur: Harmonic Analysis and Group Representations. Notices of the AMS. Gorgeous introduction to the ideas of Harish-Chandra. Available off AMS website.

I. Gelfand, M. Graev and I. Piatetskii-Shapiro: Representation Theory and Automorphic Functions. Academic Press. A classic thorough study of representations of SL2 over real, p-adic and adelic fields.

D. Bump: Automorphic Forms and Representations. Cambridge Studs. Advanced Math. 55. A good and detailed introduction to Langlands program ideas focussed on representations of GL2.

T. Bailey and A. Knapp (eds): Representation Theory and Automorphic Forms. Proc. Symp. Pure Math 61. Proceedings of an instructional conference, with a variety of great introductory articles. (In particular Schmid, Knapp and Langlands).

J. Bernstein and S. Gelbart (eds): An Introduction to the Langlands Program. Birkhauser. A very timely collection of introductions to the basic constituents of the Langlands program.

A. Borel and W. Casselman (eds): Automorphic Forms, Representations and L-functions (the Corvallis volumes). The standard source of information about the Langlands program. Available off the AMS website (see e.g. Arinkin's webpage, below).

J. Bernstein, Courses on Eisenstein Series and Representations of p-adic Groups. Beautiful expositions. Available at Dima Arinkin's Langlands page

Wee Teck Gan, Automorphic Forms and Automorphic Representations: slides for a series of five lectures given in Hangzhou, China giving an excellent overview of the basic theory (available on his web page).

Casselman, Continuous representations