MATH 210C NOTES: LIE THEORY

ARUN DEBRAY
SEPTEMBER 23, 2015

These notes were taken in Stanford's Math 210C class in Spring 2013, taught by Akshay Venkatesh. I live-TEXed them using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu.

CONTENTS

1. REPRESENTATIONS OF $\mathrm{GL}_d(\mathbb{C})$: 3/31/14

The course website is math.stanford.edu/~akshay/math210c.html. There will be homeworks about every 2 weeks, along with a take-home midterm and final. The course will assume some background knowledge about the topology of manifolds, integration of differential forms, and some familiarity with the representation theory of finite groups.

We could just start with "let G be a compact Lie group," but it's nicer to have some motivation. Let V be a finite-dimensional complex vector space. Then, what other vector spaces W can we construct "algebraically" from V? For example, we have $V \otimes V$, $\Lambda^3 V$, and so on. Note that there are no compact groups in this statement; it's purely algebra. However first we have to clarify exactly what is meant by algebraic. It's in some sense a matter of taste, as some people (but not in this class) consider the conjugate space to be algebraic, but here we want W to be functorial in V, so that in particular there is a homomorphism $\rho : \mathrm{GL}(V) \to \mathrm{GL}(W)$. But more precisely, this ρ should be an algebraic function in the following sense.

Definition 1.1. A representation $\rho : \mathrm{GL}_d(\mathbb{C}) \to \mathrm{GL}_m(\mathbb{C})$ is algebraic if the component functions $\rho(g)_{ij}$ are each algebraic functions on $\mathrm{GL}_d(\mathbb{C})$, i.e. polynomials in g_{ij} and $\det(g)^{-1}$.

The latter term appears because we need to encode the information that $\det(g) \neq 0$.

Remark. The result below will still work if one uses holomorphic functions (and representations) instead of algebraic ones. ◄

Theorem 1.2 (Weyl)**.**

(1) *Every algebraic representation of* $\mathrm{GL}(V)$ *is a direct sum of irreducible algebraic representations.*
(2) *If* $d = \dim(V)$, *then the irreducible representations of* $\mathrm{GL}(V)$ *are parameterized by* d*-tuples of integers* $n_1 \le n_2 \le \cdots \le n_d$. *The character of the irreducible representation corresponding to this* d*-tuple is*

$$\mathrm{trace}(g \in \mathrm{GL}(V)) = \frac{\det\left(x_i^{n_j+j+1}\right)}{\det\left(x_i^{j-1}\right)},$$

where $x_1, \ldots, x_d$ *are the eigenvalues of* $g \in \mathrm{GL}(V)$ *acting on* V.[1]

The proof of this theorem isn't hard, but uses Lie groups in an important way, and will draw on materials from the first part of this class.

Example 1.3.

(1) $(n_d, \ldots, n_1) = (1, 0, \ldots, 0)$ corresponds to $\mathrm{GL}(V)$ acting on V itself. For $d = 3$, the character is

$$\det \begin{bmatrix} x_1^3 & x_2^3 & x_3^3 \\ x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{bmatrix} \Big/ \det \begin{bmatrix} x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{bmatrix} = x_1 + x_2 + x_3.$$

(2) $(n_d, \ldots, n_1) = (m, 0, \ldots, 0)$, which intuitively says that the ratio of the two determinants is the sum of all monomials of degree m. This ends up being a representation of $\mathrm{GL}(V)$ on $\mathrm{Sym}^m V$. In general, each tuple comes from some functorial construction on vector spaces, but they are often non-obvious. This one will be on the first homework.
(3) $(n_d, \ldots, n_1) = (\underbrace{1, \ldots, 1}_{\ell}, 0, \ldots, 0)$ corresponds to a representation of $\mathrm{GL}(V)$ on $\bigwedge^\ell V$, i.e. the ℓ^{th} exterior power.
(4) $(n_d, \ldots, n_1) = (2, 2, 0, \ldots, 0)$ is more exotic — but a better example of the typical structure of these representations. This is a representation of $\mathrm{GL}(V)$ on

$$\left\{ x \in V^{\otimes 4} \;\middle|\; \begin{array}{l} (1\,3)x = (2\,4)x = -x \\ (1\,2)(3\,4)x = x \\ x + (2\,3\,4)x + (2\,4\,3)x = 0 \end{array} \right\}.$$

(Here, S_4 acts on $V^{\otimes 4}$ by permuting the entries, e.g. $(1\,3)(x_1 \otimes x_2 \otimes x_3 \otimes x_4) = x_3 \otimes x_2 \otimes x_1 \otimes x_4$.) These are exactly the symmetries of the Riemann curvature tensor of a Riemannian manifold! ◄

There's no great reference for this; look anywhere people talk about functors from vector spaces to themselves.

The proof of Theorem 1.2 will make several assumptions which will be justified later along in the course. Specifically, we need that:

- the representation theory of compact groups behaves exactly like that of finite groups (e.g. character orthogonality, all of the elementary structural theorems), except replacing sums with integrals, and
- an integration formula for the unitary group $\mathrm{U}(d)$, to be given more precisely later.

Definition 1.4. The unitary group is $\mathrm{U}(d) \subset \mathrm{GL}_d(\mathbb{C})$, the set of $g \in \mathrm{GL}_d(\mathbb{C})$ preserving the standard Hermitian form $\sum_{i=1}^d |z_i|^2$, i.e. so that $g\overline{g}^t = \mathrm{Id}$. This is a compact topological group, as the constraint forces individual entries in these matrices to be bounded.

[1] This is a symmetric function in the x_i.

There's no *a priori* reason to introduce the unitary group here, but it will be quite useful.

Claim. $\mathrm{U}(d)$ is Zariski-dense in $\mathrm{GL}_d(\mathbb{C})$: in other words, any algebraic function on $\mathrm{GL}_d(\mathbb{C})$ that vanishes on $\mathrm{U}(d)$ vanishes everywhere.

Proof. This will be on the first homework; it's not difficult, but involves showing said algebraic function has zero derivative. For example, when $d=1$, $\mathrm{GL}_1(\mathbb{C})\cong\mathbb{C}^*$, and $\mathrm{U}(1)$ is the unit circle, so this boils down to the fact that a holomorphic function that vanishes on the unit circle vanishes everywhere.

This result allows one to promote lots of stuff from $\mathrm{U}(d)$ to $\mathrm{GL}_d(\mathbb{C})$. In particular, if W is an algebraic representation of $\mathrm{GL}_d(\mathbb{C})$ and $W'\subset W$ is preserved by $\mathrm{U}(d)$, then W' is preserved by all of $\mathrm{GL}_d(\mathbb{C})$ (which will be explained in a moment).

Proof of Theorem 1.2, part (1). With these assumptions, we can now demonstrate the first part of the theorem: if W is a representation of $\mathrm{GL}_d(\mathbb{C})$, then split $W=\bigoplus W_i$ as $\mathrm{U}(d)$-representations (since Maschke's theorem holds for compact groups), but then, each W_i is a $\mathrm{GL}_d(\mathbb{C})$-representation too, and is irreducible because it's irreducible under the smaller group $\mathrm{U}(d)$.

Now we can go back and answer the statement given just before the proof: for $W'\subset W$ to be preserved by $\mathrm{U}(d)$ is to say that for all $u\in\mathrm{U}(d)$ and $w_1\in W'$, $u\cdot w\in W'$, or equivalently, $\langle uw_1,w_2\rangle=0$ for all $u\in\mathrm{U}(d)$, $w_1\in W'$, and $w_2\in(W')^\perp$. $\langle \underline{\quad} w_1,w_2\rangle$ is a holomorphic function on $\mathrm{GL}_d(\mathbb{C})$ that vanishes on $\mathrm{U}(d)$, and thus it vanishes everywhere, so W' is $\mathrm{GL}_d(\mathbb{C})$-invariant. ⊠

This proof used a general trick of encoding a statement that one want to generalize into a function that vanishes somewhere.

The second part of Theorem 1.2 is different; unlike for finite groups, it's possible to compute the characters just by pure thought, using the orthogonality relations but not worrying about what the representations actually are. For this part, we'll need the following facts.

- For any compact topological group, there's a preferred measure (i.e. way to integrate functions). More precisely, there's a unique measure μ such that the total measure $\mu(G)=1$ and μ is left- and right G-invariant, i.e. $\mu(S)=\mu(Sg)=\mu(gS)$ for any $g\in G$ and measurable $S\subset G$.

  A measure with total mass 1 is a probability measure, so this says there's a preferred way to talk about a random element of a compact topological group.
- The characters of a continuous representation of a topological group are orthonormal, i.e. if $V_1,V_2,\dots$ are the irreducible continuous representations of a compact topological group G, their characters $\chi_1,\chi_2,\dots$ form an orthonormal basis of the (not necessarily finite-dimensional) Hilbert space of class functions
$$\{f\in L^2(G,\mu):f(xgx^{-1})=f(g)\}.$$
- We also need an integration formula. Let
$$T=\left\{\begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n\end{pmatrix}:|z_i|=1\right\},$$
  which is a closed subgroup of $\mathrm{U}(d)$. Then, any element of $\mathrm{U}(d)$ is conjugate to an element of T (which is just a restatement of the fact that any unitary matrix can be diagonalized), so one ought to be able to tell everything about a class function from its restriction to T. Specifically, if f is a class function on $\mathrm{U}(d)$, then
$$\int_{\mathrm{U}(d)} f\,\mathrm{d}\mu=\frac{1}{n!}\int_T f\cdot|[|\Bigg]\prod_{i\le j} z_i-z_j{}^2\,\mathrm{d}\mu_T,$$
  where μ is the canonical measure discussed above for $\mathrm{U}(d)$, and μ_T is that for T. This is a special case of something we'll see again, called the Weyl integration formula.

2. Representations of $\mathrm{GL}_d(\mathbb{C})$, part II: 4/2/14

Recall that we've already proven that every algebraic (or holomorphic) representation of $\mathrm{GL}_d(\mathbb{C})$ is a sum of irreducibles, and we're trying to show that the representations are indexed by integers $m_1\le\cdots\le m_d$ with characters

$\det(x_i^{m_j+j-1})/\det(x_i^{j-1})$, where $x_1,\ldots,x_d$ are the eigenvalues of $g\in \mathrm{GL}_d(\mathbb{C})$. This was in the context of determining which vector spaces one can construct from a given one.

To prove part 2 of Theorem 1.2, we will show that the character formula holds for the unitary group $\mathrm{U}(d)\subseteq \mathrm{GL}_d(\mathbb{C})$, which is a compact subgroup, and then extend it to the whole of $\mathrm{GL}_d(\mathbb{C})$. The important feature of $\mathrm{U}(d)$ is that it's Zariski-dense, so an algebraic function on $\mathrm{GL}_d(\mathbb{C})$ that vanishes on $\mathrm{U}(d)$ must vanish everywhere.

The proof will lean on the following facts, which were presented last lecture, and will be proven formally later in the class.

- There's a unique bi-invariant (i.e. left- and right-invariant) probability measure μ on a given compact topological group, called the Haar measure. Later on in the course, we'll be able to fairly easily given an explicit formula for it.
- The characters on $\mathrm{U}(d)$ form an orthonormal basis for class functions in $\mathrm{U}(d)$ (i.e. functions that are invariant under conjugacy).
- The Weyl integration formula: if f is a continuous class function on $\mathrm{U}(d)$, then

  $$\int_{\mathrm{U}(d)} f\,\mathrm{d}\mu = \frac{1}{d!}\int_T f|\mathscr{D}|^2\,\mathrm{d}\mu_T,$$

  where T (which stands for "torus") is the diagonal subgroup of $\mathrm{U}(d)$, and $\mathscr{D}$ is given by

  $$\mathscr{D}\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_d\end{pmatrix} = \prod_{i<j}(x_i-x_j),$$

  and μ_T is the Haar measure for T (as T is also a compact topological group), given by $\mu_T=\prod \mathrm{d}\theta_i/2\pi$.

This third fact is not intuitive, but means that if one chooses a random element of $\mathrm{U}(d)$ (using the Haar measure as a probability distribution), then its eigenvalues, given by $e^{i\theta_1},\ldots,e^{i\theta_d}$ since the matrix is unitary, have a distribution function

$$\frac{1}{d!}\prod_{j<k}|e^{i\theta_j}-e^{i\theta_k}|.$$

Qualitatively speaking, this means that it's unlikely for eigenvalues to be close together. You can check this on a computer, picking a matrix and calculating the minimum distance between any two of its eigenvalues, and this result is true of other classes of matrices (e.g. symmetric matrices: those with repeated eigenvalues form a subspace of codimension greater than 1).

Proof of Theorem 1.2, part 2. Suppose V is a continuous irreducible representation of $\mathrm{U}(d)$, and let $\chi_V:\mathrm{U}(d)\to\mathbb{C}$ be its character. Then, $\chi_V\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_d\end{pmatrix}$ is given by a polynomial in $x_1,\ldots,x_n,x_1^{-1},\ldots,x_n^{-1}$. Look at it as a T-representation: since T is a compact abelian group, then the representation decomposes into a sum of irreducibles, and each irreducible has dimension 1 (since it's abelian; this is just as in the finite case).

The one-dimensional, continuous representations of T are given by

$$\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_d\end{pmatrix}\longmapsto x_1^{k_1}\cdots x_d^{k_d},\quad k_1,\ldots,k_d\in\mathbb{Z}.$$

This isn't obvious, but isn't too hard to check.

Thus, χ_V on T is a sum of these monomials. In fact,

$$\chi_V=\sum m_{\mathbf{k}}x_1^{k_1}\cdots x_d^{k_d},\quad m_{\mathbf{k}}\in\mathbb{Z}_{\geq 0}.$$

Since χ_V is polynomial in $x_1, \dots, x_d$, call this polynomial $P(x_1, \dots, x_d)$. Now, we can use orthogonality of characters:

$$\begin{aligned}\langle \chi_V, \chi_V \rangle &= \int_{\mathrm{U}(d)} |\chi_V|^2 \, \mathrm{d}\mu \\ &= \frac{1}{d!} \int_T |\chi_V \mathscr{D}|^2 \, \mathrm{d}\mu_T,\end{aligned}$$

and $\chi_V \mathscr{D}(x_1, \dots, x_d) = P(x_1, \dots, x_d)\mathscr{D}(x_1, \dots, x_d)$; $\mathscr{D}$ is anti-symmetric by its definition, and P is symmetric (because this comes from $\mathrm{U}(d)$), so we can write

$$\chi_V \mathscr{D} \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_d \end{pmatrix} = \sum m'_{\mathbf{k}} x_1^{k_1} \cdots x_d^{k_d},$$

where the $m'_{\mathbf{k}} \in \mathbb{Z}$ are anti-symmetric (and they're integers because everything else here is).

From the orthogonality of characters or Fourier series theory, the monomials $x_1^{k_1} \cdots x_d^{k_d}$ form a basis for $L^2(T)$, so, since V is irreducible, then

$$1 = \langle \chi_V, \chi_V \rangle = \frac{1}{d!} \sum_{\mathbf{k}} |m'_{\mathbf{k}}|^2.$$

Now, since $m'_{\mathbf{k}}$ is anti-symmetric in $k_1, \dots, k_d$, then it's only nonzero when the k_i are all distinct. Thus, for each $m_{(k_1,\dots,k_d)} \neq 0$, any permutation $k_i \mapsto k_{\sigma(i)}$ fixes $m_{\mathbf{k}}$ up to sign, and thus there's a unique $k_1, \dots, k_d$ up to permutation such that $m_{\mathbf{k}} \neq 0$, and thus $m_{\mathbf{k}} = \pm 1$. Thus, we can rewrite the character as

$$\chi_V = \frac{\pm \sum_{\sigma \in S_d} x_1^{k_{\sigma(1)}} \cdots x_d^{k_{\sigma(d)}} \operatorname{sign}(\sigma)}{\prod_{i<j} (x_i - x_j)}.$$

To determine the sign in the numerator, one can plug in $x_1 = 1, \dots, x_d = 1$, as $\chi_V(\mathrm{id}) > 0$.

The denominator is exactly the Vandermonde determinant

$$\prod_{i<j} (x_i - x_j) = \det \begin{bmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_d \\ \vdots & \ddots & \vdots \\ x_1^{d-1} & \dots & x_d^{d-1} \end{bmatrix},$$

and the numerator is also a determinant (the cofactor expansion formula), specifically the matrix whose ij^{th} entry is $x_j^{k_i}$. After reordering, we can assume that $k_1 < k_2 < \cdots < k_d$.[2]

So now we know that if V is an irreducible representation of $\mathrm{U}(d)$, then $\chi_V = \chi_{\mathbf{m}}$ as described in the theorem statement, for some $m_1 \le \cdots \le m_d$. We still need to show that every $\mathbf{m}$ occurs, and that these representations can be extended to $\mathrm{GL}_d(\mathbb{C})$. The proof is pretty amazing: Weyl used so little explicitly to actually compute all of the characters!

Suppose there exists an $\mathbf{m}_0$ such that $\chi_{\mathbf{m}_0}$ doesn't occur as the character of an irreducible representation of $\mathrm{U}(d)$. Then, pick any irreducible representation V, and in a similar computation to above, one can show that $\left\langle \chi_{\mathbf{m}_0}, \chi_V \right\rangle = \left\langle \chi_{\mathbf{m}_0}, \chi_{\mathbf{m}} \right\rangle$ for some $\mathbf{m} \neq \mathbf{m}_0$, and thus that it's zero. Thus, $\chi_{\mathbf{m}_0}$ is orthogonal to the space spanned by all of the irreducible characters… but they have to span the space of all class functions, so this is a contradiction.

Now, all we have to prove is the following claim:

Claim. If V is an irreducible representation of $\mathrm{U}(d)$, then it extends uniquely to $\mathrm{GL}_d(\mathbb{C})$, i.e. the homomorphism $\mathrm{U}(d) \to \mathrm{GL}(V)$ extends uniquely to an algebraic $\mathrm{GL}_d(\mathbb{C}) \to \mathrm{GL}(V)$.

A better way to state this: restriction from $\mathrm{GL}_d(\mathbb{C}) \to \mathrm{GL}(V)$ to $\mathrm{U}(d) \to \mathrm{GL}(V)$ is an equivalence of categories!

We saw that uniqueness follows because $\mathrm{U}(d)$ is Zariski-dense in $\mathrm{GL}_d(\mathbb{C})$, but for existence, we need to access the representation somehow, not just its character. It's enough to show that there's a $\mathrm{U}(d)$-representation $V' \supset V$ that extends to $\mathrm{GL}_d(\mathbb{C})$, because last time, we showed that $\mathrm{U}(d)$-invariance of a sub-$\mathrm{GL}_d(\mathbb{C})$-representation implies $\mathrm{GL}_d(\mathbb{C})$-invariance, so this implies that V lifts.

[2] In the theorem statement, we used $m_1 = k_1$, $m_2 = k_2 - 1$, $m_3 = k_3 - 2$, and so on.

Later in the course, we will explicitly write down representations of $\mathrm{U}(d)$, but for now, we'll use the following:

Definition 2.1. If G is a compact group, a translate of a character χ is a function $g \mapsto \chi(hg)$ for some $h \in G$. Then, $\mathscr{T}(\chi) \subseteq L^2(G)$ denotes the set of all translates of χ, and the inclusion is as finite-dimensional subspaces (which is a result from character theory).

Lemma 2.2. *Let G be a compact group, U an irreducible representation of G, and W a faithful representation of G. Then:*

(1) *U occurs inside $W^{\otimes a} \otimes \widetilde{W}^{\otimes b}$ for some a and b (where $\widetilde{W}$ is the dual representation).*
(2) *One can realize $U \subseteq \mathscr{T}(\chi_U)$.*

To prove (1), the goal is to approximate the character of the regular representation, which does contain all irreducible representations but might not be finite-dimensional (since G can be infinite) with that of $W^{\otimes a} \otimes \widetilde{W}^{\otimes b}$. The argument can be done with finite groups and then generalized. For (2), check for finite groups, and the same proof works here.

Armed with the above lemma, here are two ways to choose such a V':

(1) Using part (1), V occurs within some $(\mathbb{C}^d)^{\otimes a} \otimes ((\mathbb{C}^d)^*)^{\otimes b}$, which $\mathrm{GL}_d(\mathbb{C})$ acts on, extending $\mathrm{U}(d)$.
(2) Using part (2), $V \subseteq \mathscr{T}(\chi_V)$, but χ_V is a polynomial in the entries of the matrix and the inverse of the determinant and is of some bounded degree. Here, the $\mathrm{U}(d)$-action extends by considering translates on $\mathrm{GL}_d(\mathbb{C})$ instead of $\mathrm{U}(d)$ (i.e. the two spaces are isomorphic, so the action by left translation lifts). ⊠

3. Lie Groups and Lie Algebras: 4/4/14

Now, with the motivation out of the way, we can start the class properly.

A Lie group is a group G with the structure of a smooth (i.e. C^∞) manifold such that the group operations of multiplication $m : G \times G \to G$ and inversion $G \to G$ given by $x \mapsto x^{-1}$.[3]

By the end of the course, we'll classify all compact Lie groups and their representations.

Example 3.1.

(1) The basic example is $\mathrm{GL}_n(\mathbb{R})$, the group of invertible $n \times n$ real matrices. This is clearly a manifold, because it's an open subset of $M_n(\mathbb{R})$, and since matrix multiplication and inversion are polynomial functions in each entry, then they're smooth.
(2) Similarly, we have $\mathrm{GL}_n(\mathbb{C})$. It has more structure as a complex manifold (and therefore a complex Lie group), because the group operations are complex analytic. This is beyond the scope of the class, though.
(3) The orthogonal group is the group of rotations in $\mathbb{R}^n$, $\mathrm{O}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid AA^{\mathrm{T}} = \mathrm{id}_n\}$. This is a smooth submanifold of $\mathrm{GL}_n(\mathbb{R})$, because it's the preimage of the regular value id_n under the map $A \mapsto AA^{\mathrm{T}}$ from $M_n(\mathbb{R})$ to the group of symmetric matrices.[4]
(4) We'll see later in the class that if G is a Lie group and $H \subset G$ is a closed subgroup, then H is a smooth submanifold and thus acquires a Lie group structure. This provides an alternate proof that $\mathrm{O}(n)$ is a Lie group.
(5) The unitary group $\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A\overline{A}^{\mathrm{T}} = \mathrm{id}_n\}$. Again, this follows because this is the preimage of id_n, which is a regular value, or because this is a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$.
(6) In general, subgroups of $\mathrm{GL}_n(\mathbb{R})$ are a good way to obtain Lie groups, such as the Heisenberg group

$$\left\{\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}\right\} \subseteq \mathrm{GL}_3(\mathbb{R}),$$

which is diffeomorphic to $\mathbb{R}^3$, or similarly one could take the subgroup of upper triangular matrices

$$\left\{\begin{pmatrix} a & x & y \\ 0 & b & z \\ 0 & 0 & c \end{pmatrix}\right\} \subseteq \mathrm{GL}_3(\mathbb{R}).$$

[3]If one just required G to be a topological space and the maps to be continuous, then G would be a topological group.

[4]A value is said to be regular for a map if the derivative has full rank at each point in the preimage. It's a standard fact that the preimage of a regular value is a smooth submanifold.

(7) There are other Lie groups which don't occur as closed subgroups of $\mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$, but they're not as important, and we'll get to them later. ◂

Lie Algebra of a Lie Group. We want to extract an invariant by linearizing the group law near the identity. Let T be the tangent space to G at the identity e. Then, the multiplication map m is smooth, so its derivative is a map $\mathrm{d}m : T_{(e,e)}(G \times G) \to T_e G$, which can be thought of as $T \oplus T \to T$.

Unfortunately, this derivative carries no information about the group G, because the derivative sends $(x, 0) \mapsto x$ and $(0, y) \mapsto y$, so $\mathrm{d}m : (x, y) \to x + y$. This is certainly a useful fact to know, but it's not a helpful invariant. So we want something which captures more information about G.

Let's look at the map $G \times G \to G$ given by $f : (g, h) \mapsto [g, h] = ghg^{-1}h^{-1}$, their commutator. If G is abelian, this is the zero map (e.g. the Lie group $\mathbb{R}^n$). But this time, $f(e, h) = e$, so $\mathrm{d}f_{(e,e)} : T \oplus T \to T$ is now zero. In other words, elements near the identity commute to first order, because multiplication resembles addition to first order. But then, what's the quadratic term of f? The second derivatives give a quadratic function $q : T \oplus T \to T$, in that every coordinate is quadratic, and specifically a sum $x_i = \sum a^i_{jk} x_j x_k$. It's still the case that $q(0, x) = q(x, 0) = 0$, which implies it's bilinear — but $f(h, g) = f(g, h)^{-1}$, so q is a skew-symmetric bilinear form. This provides information about how two elements near the identity fail to commute.

Example 3.2. When $G = \mathrm{GL}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$, it's an open subgroup, so $T_e G \cong M_n(\mathbb{R})$. Thus, this bilinear map is $B : M_n \times M_n \to M_n$ given by $X, Y \mapsto XY - YX$. This will be the basic example of a Lie algebra.

To show the above claim and compute B, we need to compute commutators near e, so for $X, Y \in \mathrm{GL}_n(\mathbb{R})$, we want to understand $[1 + \varepsilon X, 1 + \varepsilon Y]$ for small ε. But using the Taylor series for $1/(1 + x)$, we can compute

$$\begin{aligned}
[1 + \varepsilon X, 1 + \varepsilon Y] &= (1 + \varepsilon X)(1 + \varepsilon Y)(1 + \varepsilon X)^{-1}(1 + \varepsilon Y)^{-1} \\
&= (1 + \varepsilon X + \varepsilon Y + \varepsilon^2 XY)(1 + \varepsilon Y + \varepsilon X + \varepsilon^2 YX)^{-1} \\
&= (1 + \varepsilon X + \varepsilon Y + \varepsilon^2 XY)(1 - \varepsilon Y - \varepsilon X - \varepsilon^2 YX + \varepsilon^2 (X + Y)^2 + O(\varepsilon^3)) \\
&= 1 + \varepsilon^2 (XY - YX - (X + Y)^2 + (X + Y)^2) + O(\varepsilon^3) \\
&= 1 + \varepsilon^2 (XY - YX) + O(\varepsilon^3).
\end{aligned}$$

One could state this more formally with derivatives, but this argument is easier to follow. ◂

Definition 3.3. For a Lie group G, the associated Lie algebra is $\mathfrak{g} = (T_e G, B)$, where B is the bilinear form given above (called q), usually denoted $X, Y \mapsto [X, Y]$ and called the Lie bracket.

Thus, for $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{Lie}(\mathrm{GL}_n(\mathbb{R})) = \{M_n(\mathbb{R}), [X, Y] = XY - YX\}$. There's also an abstract, axiomatic notion of a Lie algebra, which we will provide later.

The notion of a Lie algebra for a Lie group seems extremely arbitrary, at least until we get to the following theorem.

Theorem 3.4 (Lie). *$\mathfrak{g}$ determines the Lie group G locally near the identity (since the invariant is taken near the identity), in the sense that if G and H are Lie groups and $\varphi : \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ is an isomorphism that preserves the Lie bracket, then there exist open neighborhoods $U \subset H$ and $V \subset G$ of their respective identities and a diffeomorphism $\Phi : U \to V$ such that $\Phi(xy) = \Phi(x)\Phi(y)$ whenever both sides are defined, and $\mathrm{d}\Phi|_e = \varphi$.*

Note that it's possible to have $x, y \in U$ such that $\Phi(xy) \notin V$, in which case we ignore them.

This theorem states that if two Lie groups have isomorphic Lie algebras, then their group multiplication operations are the same near the identity. In fact, there's a result called the Campbell-Baker-Hausdorff formula: there exist neighborhoods $U \subset \mathfrak{g}$ of 0 and $V \subset G$ of e, a diffeomorphism $\Phi : U \to V$, and a power series $F(X, Y)$ in X, Y, and $[X, Y]$ given by

$$F(X, Y) = X + Y + \frac{[X, Y]}{2} + \frac{[X - Y, [X, Y]]}{12} + \cdots$$

such that $\Phi(F(X, Y)) = \Phi(X)\Phi(Y)$ whenever both sides are defined.

In other words, there exist local coordinates near the identity where multiplication can be written solely in terms of the Lie bracket. Thus, this single invariant locally determines the group operation!

Remark. But without this hindsight, it's completely non-obvious how to come up with the Lie bracket in the first place. However, it's possible to obtain it by carefully thinking about the Taylor expansion of $m : G \times G \to G$. (This is actually true more generally when one considers invariants given by some structure on a manifold.)

Pick a basis $e_1, \ldots, e_n$ for $T = T_e G$, and pick a system of coordinates $x_1, \ldots, x_n$ giving a local system of coordinates at the identity, i.e. an $f = (x_1, \ldots, x_n) : G \to \mathbb{R}^n$ sending $e \mapsto 0$ and choose them to be "compatible with the e_i," i.e. so that $\partial x_i|_e = e_i$.

This allows one to think of a neighborhood of $e \in G$ as a neighborhood of $0 \in \mathbb{R}^n$, and to transfer multiplication over. In particular, we can multiply things near 0, as

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = f(f^{-1}(a_1, \ldots, a_n) \cdot f^{1}(b_1, \ldots, b_n)).$$

The first-order term must be addition, because as shown before, the derivative is addition, and the second-order term is some quadratic form $Q : T \oplus T \to T$. In other words,

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n) + Q(a_1, \ldots, a_n, b_1, \ldots, b_n) + \cdots.$$

Quadratization, unlike linearization, depends on coordinates... but Q is actually bilinear, though not skew-symmetric. When one changes coordinates, one might have

$$x_i \leftarrow x_i + \sum a_{ijk} x_j x_k.$$

Then, such a coordinate change sends $Q \mapsto Q + S$ for a symmetric bilinear form S. Thus, the class of Q is well-defined in the quotient group of bilinear quadratic forms modulo symmetric bilinear forms; since every bilinear form can be decomposed into symmetric and skew-symmetric components, then this is just the group of skew-symmetric bilinear forms $T \times T \to T$.

This story works for other structures; for example, if one starts with the metric, the result is the Riemann curvature tensor! ◂

4. The Exponential Map: 4/7/14

Last time, we discussed Lie algebras: if G is a Lie group, then its tangent space T at the identity is equipped with a canonical bilinear map, the Lie bracket $[,] : T \times T \to T$, given by the quadratic term of the commutator map $g, h \mapsto ghg^{-1}h^{-1}$ (since its derivative is zero).

If two Lie groups have the same Lie algebra $(T, [,])$, then there's a smooth local identification (i.e. of some neighborhoods of the respective identities) near the identity.[5]

There is a group-theoretic proof of this theorem (involving a canonical coordinate system) and a manifold-theoretic proof. First, we will sketch the group-theoretic proof.

Proposition 4.1. *Let G be a Lie group and $T = T_e G$. For every $X \in T$, there's a unique smooth homomoprhism $\varphi_X : \mathbb{R} \to G$ such that $\varphi_X'(0) = X$. Define* $\exp(X) = \varphi_X(1)$ *giving* $\exp : T \to G$*; then, this map is smooth,* $\exp(0) = 1$*, and its derivative at* 0 *is the identity.*

This implies that exp gives a diffeomorphism from a neighborhood of $0 \in T$ to a neighborhood of $e \in G$.

Example 4.2. For $G = \mathrm{GL}_n(\mathbb{R})$, we can identify $T = M_n(\mathbb{R})$. Then, $\varphi_X(t) = e^{tX}$ for $X \in M_n(\mathbb{R})$.[6] Thus, $\exp(X) = e^X$. ◂

Remark. Note that if $G \le \mathrm{GL}_n(\mathbb{R})$ is a closed subgroup (which we will later show implies that it's also a submanifold, giving it a Lie group structure), then $\varphi_X(t) = e^{tX}$ because of the uniqueness criterion in Proposition 4.1, so $\exp(X) = e^X$ again. In particular, $e^X \in G$ when $X \in T_e G$. Thus, for all practical purposes, one can think of this exponential map as a matrix exponential. ◂

It's also convenient that with a natural choice of metric (i.e. that inducing the Haar measure), this exponential will coincide with the Riemannian exponential.

Proof sketch of Proposition 4.1. We want $\varphi_X(t) = \varphi_X(t/N)^N$, and if N is large, this is near the identity (so that the derivative is about 0). This is enough for a unique definition: if $\varphi : \mathbb{R} \to G$ is such that $\varphi'(0) = X$, then $\lim_{N\to\infty} \varphi(t/N)^N$ exists and is independent of φ, so this must be $\varphi_X(t)$, which will imply existence and uniqueness.

[5] There's a cleaner statement of this, also known as Lie's theorem, involving covering spaces, which we'll cover later on in this course.

[6] Here, the matrix exponential is defined via the usual power series or as

$$e^X = \lim_{N\to\infty}\left(1 + \frac{X}{N}\right)^N,$$

so that $e^A e^B = e^{A+B}$. The second definition is useful because it makes the linear term easy to extract.

It's pretty intuitive, but the details are rather annoying to write out, which is why people don't often prove Lie's theorem in this way. First off, it's easier to take $\lim_{N\to\infty} \varphi(t/2^N)^{2^N}$, which will be the same.

Suppose φ_1 and φ_2 are such that $\varphi_1'(0) = \varphi_2'(0) = X$, and such that $\varphi_1(0) = \varphi_2(0) = e$. Write $f(t) = \varphi_2^{-1}\varphi_1$, so that $\varphi_1(t) = \varphi_2(t)f(t)$. Then, $f'(0) = 0$ (since multiplication near e looks like addition), so if we look at the Taylor series, $f(1/N)$ is quadratically close to e (i.e. the distance is $\ll 1/N^2$, measured in some coordinates near the identity).

Then,

$$\begin{aligned}\varphi_1\Big(\frac{1}{M}\Big)^M &= \overbrace{\varphi_2\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big)\varphi_2\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big)\cdots\varphi_2\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big)}^{M\text{ times}}\\ &= \varphi_2\Big(\frac{1}{M}\Big)^2\Big(\varphi_2^{-1}\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big)\varphi_2\Big(\frac{1}{M}\Big)\Big)\varphi_2\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big)\cdots\varphi_2\Big(\frac{1}{M}\Big)f\Big(\frac{1}{M}\Big).\end{aligned}$$

Applying this many times,

$$= \varphi_2\Big(\frac{1}{M}\Big)^M \underbrace{\varphi_2\Big(\frac{1}{M}\Big)^{-(M-1)} f\Big(\frac{1}{M}\Big)\varphi_2\Big(\frac{1}{M}\Big)^{M-1}\varphi_2\Big(\frac{1}{M}\Big)^{-(M-2)} f\Big(\frac{1}{M}\Big)\varphi_2\Big(\frac{1}{M}\Big)^{M-2}\cdots f\Big(\frac{1}{M}\Big)}_{(*)}.$$

Then, the claim is that $(*)$ is small, i.e. that its distance to the identity is at most some constant over M. This is true because it's a product of M terms, each of which is on the order of $1/M^2$ distance form the identity. This argument can be made more precise by writing out Taylor's theorem a lot, but this is messy. Thus,

$$\varphi_1\Big(\frac{1}{M}\Big)^M = \varphi_2\Big(\frac{1}{M}\Big)^M h,$$

for a constant h depending on the coordinate chart such that $\operatorname{dist}(h, e)$ is less than a constant times $1/M$. The actual derivation isn't pretty, but it works, and only uses the group structure!

Now, apply this to $\varphi_2 = \varphi_1(t/2)^2$, so that $\varphi_2' = \varphi_1'$, and with $M = 2N$. This implies that

$$\varphi_1\Big(\frac{t}{2^N}\Big)^{2N} = \varphi_1\Big(\frac{t}{2^{N+1}}\Big)^{2N+1} h,$$

where h has the same restrictions as before. Then, the sequence must converge, because it's moving a distance of $1/2N$ for each N. (Of course, there's something to demonstrate here.) But the point is, the limit $\lim_N \varphi(t/2^N)^{2^N}$ exists and is independent of φ, and at this point it's easy to check that φ must be a homomorphism, though one also must show that exp is smooth in X. ⊠

Next, we want to show that there's a universal expression for the product in these exponential coordinate, i.e.

$$\exp(X)\exp(Y) = \exp(F(X,Y)) = \exp\Big(X + Y + \frac{[X,Y]}{2} + \frac{[X-Y,[X,Y]]}{12} + \cdots\Big),$$

given by the Baker-Campbell-Hausdorff power series, completely independent of G. The terms are a mess: nobody really needs to write them down, and it's much more important that such a formula exists. Then, there will be a neighborhood of 0 in T such that F converges and equality holds for $X, Y \in T$. This induces the required isomorphism of Lie groups — it's pretty fantastic that the single bilinear form (in a finite-dimensional space) completely classifies everything!

Example 4.3. Once again consider $G = \mathrm{GL}_n(\mathbb{R})$, to see that this is not trivial. Now, the statement means that $e^X e^Y = e^Z$ for some Z. For $M \in \mathrm{GL}_n(\mathbb{R})$ near the identity, the logarithm is defined by its power series:

$$\log(M) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}(M-1)^i,$$

so that $\log(e^X) = X$.

Now one can compute $\log(e^X e^Y)$, which can be done by multiplying out the respective power series:

$$\begin{aligned}\log(e^X e^Y) &= \log\left(\left(1+X+\frac{X^2}{2!}+\cdots\right)\left(1+Y+\frac{Y^2}{2!}+\cdots\right)\right)\\ &= \log\left(1+X+Y+XY+\frac{X^2}{2}+\frac{Y^2}{2}+\cdots\right)\\ &= X+Y+XY+\frac{X^2}{2}+\frac{Y^2}{2}-\frac{(X+Y)^2}{2}+\cdots\\ &= X+Y+\frac{XY-YX}{2}+\cdots\end{aligned}$$

There's a lot of ugly terms here if one goes forward, but the content of the theorem is that everything can be expressed in terms of the Lie bracket, and no multiplication within the group. ◀

Dynkin gave a combinatorial formula for this, replacing each coefficient with the corresponding Lie brackets such that the overall sum is the same, e.g. $YYX \mapsto [Y,[Y,X]]/3$, where the denominator varies with each term.

So, why is all of this believable? The things one needs to fill in are estimating errors to ensure they don't blow up. We can quantify this error, though there's no reason to go into huge detail. We had $\exp(X) = \exp(X/N)^N$ for some large N, so let $g = \exp(X/N)$ and $h = \exp(Y/N)$. Thus, $\exp(X)\exp(Y) = g^N h^N$, and $\exp(X+Y) = \exp((X+Y)/N)^N \approx (gh)^N$ up to terms of quadratic order. The computation this formula gives is the error term going from $g^N h^N$ to $(gh)^N$, and the error term is several factors of $(g^{-1}h^{-1}gh)$, which near the identity is determined by the Lie bracket. The actual computation, though, is rather ugly.

5. The Baker-Campbell-Hausdorff Formula: 4/9/14

Recall that last time, we let G be a Lie group and $T = T_e G$, along with a Lie bracket $T \times T \to T$, which indicates how the commutator behaves near the origin. Then, we defined the exponential map $\exp : T \to G$ such that for any $X \in T$, $\exp(tX)$ is a smooth homomorphism $\mathbb{R} \to G$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\exp(tX))\bigg|_{t=0} = X.$$

The goal of the Baker-Campbell-Hausdorff formula is to write $\exp(X)\exp(Y) = \exp(F(X,Y))$, where F is a power series given by the Lie bracket — the point is really that such a formula exists (and Dynkin gives a much cleaner result once a formula is shown to exist). Part of the statement is checking that all of the relevant error terms go to 0, and that the series converges. These are not difficult, but will be omitted.

Without further computation, one can deduce that F must be a smooth function, and by the group laws, $F(-X,-Y) = -F(X,Y)$, because $\exp(-X) = \exp(X)^{-1}$ (since it's a homomorphism), and that $F(X,0) = X$ and $F(0,Y) = Y$. Thus, $F(X,Y) = X+Y+B(X,Y)+\cdots$, where the remaining terms are at least cubic in X and Y, and $B : T \times T$ is skew-symmetric. Then, by computing the commutator from the Lie bracket, $B(X,Y) = [X,Y]/2$. Thus, the point is to show that the higher-order terms are determined.

Example 5.1.

(1) When $G = \mathrm{GL}_n(\mathbb{R})$, then $T = M_n(\mathbb{R})$ and $\exp = e^X$.

(2) For the orthogonal group $\mathrm{O}(n)$, this is a smooth submanifold, so (as in the homework) its tangent space is $T = \{X \in M_n(\mathbb{R}) \mid X + X^{\mathrm{T}} = 0\}$, and by the uniqueness of the exponential map, $\exp(X) = e^X$ still. Thus, the exponential of a skew-symmetric matrix is orthogonal (which can be checked by other means… but this is still pretty neat).

(3) If $G = \mathrm{U}(n)$, then $T = \{X \in M_n(\mathbb{R}) \mid X + \overline{X}^{\mathrm{T}} = 0\}$ (i.e. the space of skew-Hermitian matrices), and $\exp(X) = e^X$ again (for the same reason: its uniqueness on $\mathrm{GL}_n(\mathbb{R})$ implies it's the same on its subgroups), so the exponential of a skew-Hermitian matrix is unitary. ◀

Functoriality. Suppose $\varphi : G \to G'$ is a homomorphism of Lie groups, i.e. a smooth group homomorphism. Then, $\mathrm{d}\varphi : T \to T'$ (where $T = T_e G$ and $T' = T_e G'$, and the derivative is understood to be at the identity) respects Lie brackets, and is thus a homomorphism of Lie algebras; that is, for all $X, Y \in T$, $[\mathrm{d}\varphi(X), \mathrm{d}\varphi(Y)] = \mathrm{d}\varphi[X,Y]$, because $\varphi(xyx^{-1}y^{-1}) = \varphi(x)\varphi(y)\varphi(x)^{-1}\varphi(y)^{-1}$, so the structure follows over. Thus, constructing the tangent space is functorial. Moreover, $\varphi(\exp X) = \exp(\mathrm{d}\varphi(X))$, because $t \mapsto \varphi(\exp(tX))$ and $t \mapsto \exp(t\,\mathrm{d}\varphi(X))$ are both homomorphisms $\mathbb{R} \to G$ with the same derivative. That is, the following diagram commutes.

$$\begin{tikzcd} T \arrow[r, "\mathrm{d}\varphi"] \arrow[d, "\exp"'] & T' \arrow[d, "\exp"] \\ G \arrow[r, "\varphi"] & G' \end{tikzcd}$$

Conjugation and the Adjoint. For $x \in G$, $g \mapsto xgx^{-1}$ is a smooth isomorphism $G \to G$, so one can define $\mathrm{Ad} : T \to T$ to be its derivative.

Example 5.2. Suppose $G = \mathrm{GL}_n(\mathbb{R})$, $Y \in M_n(\mathbb{R})$, and $x \in G$. Then, $\mathrm{Ad}(x)Y = xYx^{-1}$ (there is something to show here; you have to compute a derivative). Then, the same formula holds for subgroups of $\mathrm{GL}_n(\mathbb{R})$. $\blacktriangleleft$

In other words, a Lie group acts on its own tangent space, and Ad is just multiplication near the identity. Since exp is compatible with homomorphisms, as discussed above, then $\exp(\mathrm{Ad}(x)Y) = x\exp(Y)x^{-1}$.

Now, there's a map $x \mapsto \mathrm{Ad}(x)$, going from $G \to \mathrm{GL}(T)$, since $\mathrm{Ad}(x)$ is always an isomorphism (which is isomorphic to $\mathrm{GL}_d(\mathbb{R})$ for some $d = \dim(T)$), so differentiating Ad induces a map $\mathrm{ad} : T \to \mathrm{Lie}(\mathrm{GL}(T)) = \mathrm{End}(T)$. This last equivalence is because for any vector space V, $T_e\mathrm{GL}(V) = \mathrm{End}(V)$, so, once one chooses coordinates, $T_e\mathrm{GL}_n(\mathbb{R}) \cong M_n(\mathbb{R})$.

Proposition 5.3. $\mathrm{ad}(X)$ *is differentiated conjugation:* $\mathrm{ad}(X)(Y) = [X, Y]$.

This makes $\mathrm{ad}(X)$ seem like elaborate notation, but it will simplify things. For example, when $G = \mathrm{GL}_n(\mathbb{R})$, then $\mathrm{ad}(X)$ is the linear transformation $Y \mapsto XY - YX$.

Proof of Proposition 5.3. The proof boils down to unraveling the definition. Differentiating $\mathrm{Ad}(x)$ involves moving in G on a curve in the direction x,

$$\begin{aligned} \mathrm{ad}(X)(Y) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}(e^{tX})(Y) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} (\exp(tX)\exp(sY)\exp(-tX)) \\ &= \left.\frac{\partial^2}{\partial t \partial s}\right|_{t,s=0} (\exp(tX)\exp(sY)\exp(-tX)). \end{aligned}$$

In the coordinates given by exp, multiplication is given by $X, Y \mapsto X + Y + [X, Y]/2 + \cdots$, so plugging in, we can extract the quadratic terms:

$$\exp(tX)\exp(sY)\exp(-tX) = \exp(sY + ts[X, Y] + \text{higher-order-terms}\ldots),$$

so the quadratic term is $[X, Y]$. $\boxtimes$

Since the Lie bracket pins down term up to quadratic order, lower-order terms can be given in terms of it (though higher-order terms require more trickery).

It will also be useful to have that $g\exp(Y)g^{-1} = \exp(\mathrm{Ad}(g)Y)$ and, since homomorphisms are compatible with the exponential map, then $\mathrm{Ad}(\exp(X)) = \exp(\mathrm{ad}(X))$.

Proving the Formula. For large N, we want $\exp(X) = \exp(X/N)^N$, so let $g = \exp(X/N)$ and $h = \exp(Y/N)$, and now we want to compare $g^N h^N$ with $(gh)^n \approx \exp(X + Y)$. Intuitively, there are a bunch of commutators, which are given in terms of the Lie bracket, but it's painful to write them out directly.

First, we'll compute $\exp(X)\exp(Y)$ to first order in Y (i.e., assuming that Y is small; there are infinitely many first-order terms in the formula, but within more and more iterated Lie brackets). Then,

$$\exp(X + Y) = \exp\left(\frac{X + Y}{N}\right)^N \approx \left(\exp\left(\frac{X}{N}\right)\exp\left(\frac{Y}{N}\right)\right)^N \tag{5.4}$$

with error quadratic in N, so that (5.4) is true as $N \to \infty$, in the limit. Then, take some commutators:

$$\begin{aligned}
&= \exp\left(\frac{X}{N}\right)\exp\left(\frac{Y}{N}\right)\cdots\exp\left(\frac{X}{N}\right)\exp\left(\frac{Y}{N}\right)\\
&= \left(\exp\left(\frac{X}{N}\right)\exp\left(\frac{Y}{N}\right)\exp\left(\frac{X}{N}\right)\right)\left(\exp\left(\frac{2X}{N}\right)\exp\left(\frac{Y}{N}\right)\exp\left(\frac{2X}{N}\right)\right)\cdots\left(\exp X\exp\left(\frac{Y}{N}\right)\exp X\right)\exp X\\
&= \exp\left(\mathrm{Ad}\left(\exp\left(\frac{X}{N}\right)\right)\frac{Y}{N}\right)\exp\left(\mathrm{Ad}\left(\exp\left(\frac{2X}{N}\right)\right)\frac{Y}{N}\right)\cdots\exp\left(\mathrm{Ad}(\exp(X))\frac{Y}{N}\right)\exp X.
\end{aligned}$$

Here, we use the fact that Y is small to put these together:

$$\approx \exp\left(\mathrm{Ad}\left(\exp\left(\frac{X}{N}\right)\right)\frac{Y}{N} + \mathrm{Ad}\left(\exp\left(\frac{2X}{N}\right)\frac{Y}{N}\right) + \cdots + \mathrm{Ad}\left(\exp(X)\frac{X}{N}\right)\right)\exp X,$$

which is true up to an error of size $||Y||^2$, i.e. to first-order in Y. (Specifically, if $||\cdot||$ is some norm on the tangent space, then $\exp(X+Y) = \exp(\mathrm{Error})\exp X\exp Y$, where $|\mathrm{Error}| \le c_X||Y||^2$, where c_X is constant in X.) Then, as $N \to \infty$, this can be replaced by an integral:

$$\begin{aligned}
&= \exp\left(\int_0^1 \mathrm{Ad}(\exp(tX))Y\,\mathrm{d}t\right)\exp(X)\\
&= \exp\left(\int_0^1 \exp(t\,\mathrm{ad}(X))Y\,\mathrm{d}t\right)\exp(X).
\end{aligned}$$

Here, $\mathrm{ad}(X) \in \mathrm{End}(T)$, so $\exp(t\,\mathrm{ad}(X)) \in \mathrm{GL}(T)$ is the usual matrix exponential.

$$= \exp\left(\frac{e^{\mathrm{ad}X} - 1}{\mathrm{ad}X}Y\right)\exp(X). \tag{5.5}$$

This integral means the power series integrated term-by-term, so for $M \in M_n(\mathbb{R})$,

$$\frac{e^M - 1}{M} = \sum_{i=1}^{\infty}\frac{1}{i!}M^{i-1}.$$

In summary, for small Y,

$$\exp(X+Y) = \exp\left(\frac{e^{\mathrm{ad}(X)} - 1}{\mathrm{ad}(X)}Y\right)\exp(X)E, \tag{5.6}$$

where E is the error, at distance less than a constant in X times $||Y||^2$ from the identity. This is kind of ghastly as a result, but the proof technique is most important, rather than the specific result.

More generally, let $Z(t)$ be such that $\exp(Z(t))) = \exp(X)\exp(Y)$, restricting to a neighborhood in which exp is a diffeomorphism. Then, $Z(0) = 0$, and we want to find $Z(1)$, which we'll derive from (5.5); then, its solution will end up only depending on iterated Lie brackets.

6. A Manifold-Theoretic Perspective: 4/11/14

First, we will finish the proof sketch of the Baker-Campbell-Hausdorff formula. The goal is to write down a formula for multiplying exponentials of a Lie group: $\exp(X)\exp(Y) = \exp(F(X,Y))$. last time, we showed (5.6), where the error term is quadratic in Y (or more precisely, its distance from the identity). Recall also that $\mathrm{ad}X : T \to T$ sends $Z \mapsto [X,Z]$.

If you're unhappy with this error term formulation, an alternate way of saying it is that

$$\left.\frac{\mathrm{d}}{\mathrm{d}Y}\right|_{Y=0}(\exp(X+Y)\exp(-X)) = \frac{\exp(\mathrm{ad}X) - 1}{\mathrm{ad}X},$$

which is a map $T \to T$. That is, (5.6) computes the derivative of exp. Thus,

$$\frac{\exp(\mathrm{ad}X) - 1}{\mathrm{ad}X}(Y) = \sum_{i\ge 1}\frac{(\mathrm{ad}X)^{i-1}}{i!}Y = Y + \frac{[X,Y]}{2} + \frac{[X,[X,Y]]}{6} + \cdots,$$

which will be all the terms in the final formula that are linear in Y.

Now, when Y isn't necessarily small, one can use (5.6) to get to the full formula.

Proof. The power series is $1+\operatorname{ad}(X)+\cdots$, i.e. a small perturbation of the identity map, so $\exp(\operatorname{ad}(X)-1)/\operatorname{ad}(X)$ is invertible for X (as a transformation or a matrix) in a neighborhood of 0. Call its inverse $\operatorname{ad}(X)/(\exp(\operatorname{ad}(X))-1)$, so when X is in this neighborhood,

$$\exp\left(X+\frac{\operatorname{ad}X}{e^{\operatorname{ad}x}-1}Y\right)=\exp(Y)\exp(X)(\text{error}), \tag{6.1}$$

where the error is still quadratic in Y. Similarly, if one commutes things in the opposite order, the X and Y on the opposite side are switched, and the error is still quadratic in Y.

Let $\log : T \to G$ sending $0 \to e$ be an inverse to exp on an open neighborhood of 0. Then, let $Z(t) = \log(\exp(tX)\exp(tY))$, so that $Z(0)=0$ and the goal is to compute $Z(1)$. Then,

$$\begin{aligned}\exp(Z(t)+\varepsilon) &= \exp((t+\varepsilon)X)\exp((t+\varepsilon)Y) = \exp(\varepsilon X)\exp(tX)\exp(tY)\exp(\varepsilon Y)\\ &= \exp(\varepsilon X)\exp(Z(t))\exp(\varepsilon Y).\end{aligned}$$

When ε is small, (6.1) applies to the right-hand side, so to first order,

$$=\exp\left(Z(t)+\frac{\operatorname{ad}Z(t)}{e^{\operatorname{ad}Z(t)}-1}(\varepsilon X)=\frac{(-\operatorname{ad}Z(t))}{e^{-\operatorname{ad}Z(t)}-1}(\varepsilon Y)\right), \tag{6.2}$$

i.e.

$$\frac{\mathrm{d}Z}{\mathrm{d}t}=\frac{\operatorname{ad}Z(t)}{e^{\operatorname{ad}Z(t)}-1}X-\frac{\operatorname{ad}Z(t)}{e^{\operatorname{ad}Z(t)}-1}Y.$$

But these are just big piles of iterated commutators, so write $Z(t)=\sum t^n Z_n$ and solve, so that Z_n is just a linear combination of iterated Lie brackets. This is totally unhelpful for actually finding the formula, but when X and Y are small, this converges, so the Baker-Campbell-Hausdorff formula does in fact exist, and is given by $Z(1)$. ⊠

The final formula is too complicated to be useful, but its existence is very helpful, as is the computation (6.2).

Approach via Manifolds. Following the book a little more closely, we'll see a manifold-theoretic approach to understanding the exponential map and Lie's theorem, as well as a very important result about closed subgroups. From now on, let $\mathfrak{g}$ denote the Lie algebra of the Lie group G.

Recall that a vector field $\mathscr{X}$ is a smooth assignment $\mathscr{X} : g \mapsto \mathscr{X}(g) \in T_g G$. Also, let left-multiplication by a $g \in G$ be denoted $l_g : G \to G$, sending $h \mapsto gh$. Since G is a Lie group, this is a smooth isomorphism.

Definition 6.3. A vector field $\mathscr{X}$ is left-invariant if for all $g,h \in G$, $\mathscr{X}(hg)=(Dl_h)\mathscr{X}(g)$.

If G is a Lie group, any $X \in \mathfrak{g}$ gives a left-invariant vector field on G, and in fact there's a unique left-invariant vector field $\mathscr{X}$ such that $\mathscr{X}(e)=X$. Thus, there's an isomorphism of vector spaces $\Psi : \mathfrak{g} \to \{\text{left-invariant } G\text{-vector fields}\}$ sending $X \mapsto \mathscr{X}$ as constructed above, its unique left-invariant extension.

In fact, the textbook defines the Lie algebra of a given Lie group as this space of left-invariant vector fields, with the vector field bracket $[\mathscr{X},\mathscr{Y}]$ given as follows: a vector field $\mathscr{X}$ acts on a function f as a derivation, i.e. $\mathscr{X}f$ is the derivative of f in the direction $\mathscr{X}$. Then, the vector bracket measures how much these fail to commute: $[\mathscr{X},\mathscr{Y}]f=(\mathscr{X}\mathscr{Y}-\mathscr{Y}\mathscr{X})f$. Under the isometry Ψ above, the Lie bracket as we defined it on $\mathfrak{g}$ is sent to this bracket, so $[X,Y]\mapsto[\mathscr{X},\mathscr{Y}]$.

Geometrically, let $\operatorname{flow}_{\mathscr{X}}(t)$ be the flow along $\mathscr{X}$ for time t. Then,

$$\operatorname{flow}_{-\mathscr{Y}}(\varepsilon)\operatorname{flow}_{-\mathscr{X}}(-\varepsilon)\operatorname{flow}_{\mathscr{Y}}(\varepsilon)\operatorname{flow}_{\mathscr{X}}(\varepsilon)\approx\operatorname{flow}_{[\mathscr{X},\mathscr{Y}]}(\varepsilon^2),$$

and as $\varepsilon \to 0$, this gives another definition.

Recall that an integral curve on a manifold M is a curve $f : \mathbb{R} \to M$ such that $\frac{\mathrm{d}f}{\mathrm{d}t}=\mathscr{X}(f(t))$ for some $\mathscr{X}$. The theorem on the existence of ODEs implies that integral curves exist and are locally unique.

Proposition 6.4. *Let $\mathscr{X}$ be a left-invariant vector field associated as above to an $X \in \mathfrak{g}$. Then, $t \mapsto \exp(tX)$ is the integral curve of $\mathscr{X}$ passing through e.*

This can actually be taken as the definition for the exponential map,[7] and makes it that much more obvious that it's smooth.

[7]As yet, it's only defined locally, but one can use the fact that it's a homomorphism to extend it to all of $\mathbb{R}$.

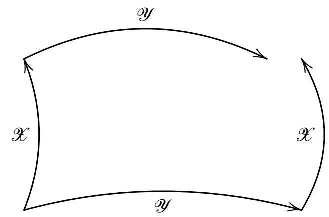

FIGURE 1. Geometric intuition behind the noncommutativity of the Lie bracket: the gap in the upper-right corner is $[\mathscr{X},\mathscr{Y}] = \mathscr{X}\mathscr{Y} - \mathscr{Y}\mathscr{X}$.

Proof of Proposition 6.4. By definition, $\frac{\mathrm{d}}{\mathrm{d}t}(\exp(X))\Big|_{t=0} = X$, so since it's a homomorphism, $\exp((s+t)X) = \exp(sX) + \exp(tX)$, so

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \exp((s+t)X) = (l_{\exp(sX)})X = \mathscr{X}(\exp(sX)). \qquad \boxtimes$$

From this point of view,[8] one can check that $[,]$ satisfies the Jacobi identity

$$[X,[Y,Z]] + [Y,[X,Z]] + [Z,[X,Y]] = 0, \tag{6.5a}$$

or, a little more nicely, using skew-symmetry alone, that

$$\mathrm{ad}(X)[Y,Z] = [Y,\mathrm{ad}(X)Z] + [\mathrm{ad}(X)Y,Z], \tag{6.5b}$$

i.e. the ad operation becomes like a derivative for the Lie bracket, in the sense of having a Leibniz rule.

Proof of (6.5a) *and* (6.5b). For (6.5a), one can expand via $[\mathscr{X},\mathscr{Y}] = \mathscr{X}\mathscr{Y} - \mathscr{Y}\mathscr{X}$. Alternatively, one could prove it via (6.5b): start with the fact that a homomorphism of Lie groups preserves the Lie bracket, so that

$$[\mathrm{Ad}(g)Y, \mathrm{Ad}(g)Z] = \mathrm{Ad}(g)[Y,Z],$$

and then differentiate once. $\boxtimes$

This motivates an alternate, abstract definition for Lie algebras.

Definition 6.6. A Lie algebra over $\mathbb{R}$ is a real[9] vector space V together with a skew-symmetric bilinear form $[,] : V \times V \to V$ satisfying the Jacobi identity.

In this sense, this discussion shows that every Lie group gives rise to a Lie algebra (in the abstract sense). It's also true that every abstract Lie algebra gives rise to a Lie group, but this ends up being less useful.

The point of the Baker-Campbell-Hausdorff formula is that a Lie algebra gives rise to a Lie group locally. There's also a manifold-theoretic version of this proof, though it doesn't lead to an explicit formula (which is OK, since we care more about its existence). Essentially, the category of real Lie algebras is equivalent to the category of simply connected Lie groups (since homomorphisms can be locally extended), which will be discussed further in a later lecture.

The proof sketch begins as follows: suppose G and G' are Lie groups such that $\varphi : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}'$, so that the goal is to locally produce a $\Phi : G \xrightarrow{\sim} G'$, meaning that $\Phi(\exp X) = \exp(\varphi(X))$. We had obtained a homomorphism $\mathbb{R} \to G$ via an integral curve earlier, and want to make this more general.

Definition 6.7. Suppose M is a manifold of dimension d. Then, $\mathscr{L} \subseteq TM$ is a subbundle of the tangent bundle of dimension k if for all $x \in M$, there is a smoothly varying family of k-dimensional vector spaces $\mathscr{L}_x \subseteq T_xM$.

When $k = 1$, a 1-dimensional subbundle is akin to a vector field, but without keeping track of magnitude. Thus, as was determined for vector fields in Proposition 6.4, one might want to "integrate" subbundles into k-dimensional submanifolds, i.e. find a k-dimensional submanifold $N \subseteq M$ such that $TM = \mathscr{L}$. That such a manifold exists means that there exists a smooth $\Phi : (\varepsilon,\varepsilon)^k \to M$, such that $D\Phi(a) = \mathscr{L}_{\Phi(a)}$, i.e. $\mathrm{Im}(\Phi)$ is a submanifold of M and its tangent space at the image of a point a is $\mathscr{L}_{\Phi(a)}$.

[8]There is a group-theoretic proof of the Jacobi identity, but it involves more fiddling around with commutators.

[9]This definition can be made over any field.

Theorem 6.8 (Frobenius)**.** *Suppose M is a d-dimensional manifold and $\mathscr{L} \subseteq TM$ is a k-subbundle. Then, letting $[\mathscr{L},\mathscr{L}] = \{[\mathscr{X},\mathscr{Y}] \mid \mathscr{X},\mathscr{Y} \subseteq \mathscr{L}\}$, there exists an integral manifold for $\mathscr{L}$ iff $[\mathscr{L},\mathscr{L}] = \mathscr{L}$.*

The proof will be given next lecture.

7. Frobenius' Theorem and Closed Subgroups: 4/14/14

Last time, we discussed Frobenius' theorem, as a means to approach Lie's theorem from a manifold-theoretic viewpoint. If G is a Lie group and $\mathfrak{g}$ its Lie algebra, then an $X \in \mathfrak{g}$ can be identified with a left-invariant vector field $\mathscr{X}$ such that $\exp(tX)$ is the integral curve for $\mathscr{X}$ through e. The end goal is to show that if G and G' have isomorphic Lie algebras, then $G \cong G'$ locally.

The proof will use Frobenius' theorem, which was given last time, and states that if M is a d-dimensional manifold and $\mathscr{L} \subseteq TM$ is a k-dimensional subbundle (i.e. for each x, $\mathscr{L}(x) \subseteq T_xM$ is a k-dimensional subspace), then if $[\mathscr{L},\mathscr{L}] = \mathscr{L}$,[10] then $\mathscr{L}$ is locally integrable, i.e. there is a local coordinate chart $U \to \mathbb{R}^d$ sending $\mathscr{L} \to \operatorname{Span}\{\partial_1,\dots,\partial_k\}$ (here these mean the first k coordinates). In other words, there's a coordinate system in which $\mathscr{L}$ looks flat.[11]

Proof of Theorem 6.8. Near some $x \in M$, pick some vector fields $X_1,\dots,X_k$ spanning $\mathscr{L}$, so that if y is in a neighborhood of x, then $\mathscr{L}(y) = \langle X_1(y),\dots,X_k(y)\rangle$. That this can be done follows from the definition of a subbundle. Then, pick some coordinate system $M \to \mathbb{R}^d$ near x, and write

$$\begin{aligned} X_1 &= a_{11}\partial_{x_1} + a_{12}\partial_{x_2} + \cdots + a_{1d}\partial_{x_d} \\ &\vdots \\ X_k &= a_{k1}\partial_{x_1} + a_{k2}\partial_{x_2} + \cdots + a_{kd}\partial_{x_d}. \end{aligned}$$

Since the X_k span a k-dimensional space, then some $k \times k$ minor of the matrix (a_{ij}) is nondegenerate, so without loss of generality, assume it's $\partial_{x_1},\dots,\partial_{x_k}$ (since if not, one can shuffle some coordinates around). Thus, after some row and column operations, it becomes diagonal, so

$$\begin{aligned} X_1 &= \partial_1 + (b_{1,k+1}\partial_{k+1} + \cdots + b_{1,d}\partial_d) \\ X_2 &= \partial_2 + (b_{2,k+1}\partial_{k+1} + \cdots + b_{2,d}\partial_d), \end{aligned}$$

and so on. These still span $\mathscr{L}$ locally, so now we can use the fact that $\mathscr{L}$ is closed under Lie bracket. Using the fact that

$$\begin{aligned} [\partial_1, f\partial_m] &= \partial_1(f\partial_m) - f\partial_m\partial_1 \\ &= \frac{\partial f}{\partial x_1}\partial_m + f\partial_1\partial_m - f\partial_1\partial_m \\ &= \frac{\partial f}{\partial x_1}\partial_m, \end{aligned}$$

$[X_1,X_2]$ must only have cross terms including ∂_ℓ for $\ell > k$. However, since these must lie in $\mathscr{L}$, then they can be written as a linear combination of the $\partial_1,\dots,\partial_k$, and thus the cross terms go to zero. Thus, $[X_i,X_j] = 0$, so $\mathscr{L}$ is spanned by commuting vector fields.

Thus, their flows also commute, so one may define the coordinate chart

$$(a_1,\dots,a_k,b_1,\dots,b_{d-k}) \longmapsto (\operatorname{flow}_{X_1}(a_1),\dots,\operatorname{flow}_{X_k}(a_k))(0,\dots,0,b_1,\dots,b_k).$$

With respect to this coordinate chart, $\mathscr{L} = \operatorname{Span}(\partial_1,\dots,\partial_k)$. ⊠

This last step can be done iff $\mathscr{L}$ is closed under Lie bracket.

Returning to Lie groups, supposing one has a $\varphi : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}'$, it will be possible to construct a $\Phi : G \xrightarrow{\sim} G'$ by looking at graphs. We already know it must obey $\Phi(\exp X) = \exp(\varphi(X))$, but need to show that it preserves multiplication.

Instead of producing Φ explicitly, one can provide its graph $\{(g,\Phi(g))\} \subset G \times G'$.

[10] This means that for all $\mathscr{X},\mathscr{Y} \in \mathscr{L}$, $[\mathscr{X},\mathscr{Y}] \in \mathscr{L}$.

[11] This is analogous to the rectification theorem for vector fields in the theory of ODEs.

Proposition 7.1. *If G is a Lie group and $\mathfrak{g}$ its Lie algebra, then if $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra (i.e. a subspace closed under the Lie bracket), then there exists a Lie group H with Lie algebra $\mathfrak{h}$, and a smooth homomorphism $f : H \to G$ inducing $\mathfrak{h} \to \mathfrak{g}$.*

Example 7.2. Let $G = \mathbb{R}^2/\mathbb{Z}^2$, so that $\mathfrak{g} = \mathbb{R}^2$. This isn't a very interesting Lie algebra, since it's commutative, so $[,] = 0$. Thus, any subspace is a Lie subalgebra, such as $\mathfrak{h} = \{(x, ax)\}$ for some given a. If $a \in \mathbb{Q}$, then there does exist a closed subgroup $H \subseteq G$ with $\mathrm{Lie}(H) = \mathfrak{h}$, but if not, then such a subgroup would be dense in G, which is the reasoning behind the seemingly clumsy wording of the fact. ◀

Proof of Proposition 7.1. For the scope of this proof, given a Lie group G and its Lie algebra $\mathfrak{g}$, let $\mathscr{X} = \mathrm{L}(X)$ denote the unique left-invariant vector field on G induced by an $X \in \mathfrak{g}$, as discussed before.

Let $\underline{\mathfrak{h}} \subset TG$ be the subbundle defined by $\{\mathrm{L}(X) : X \in \mathfrak{h}\}$. Then, $[\underline{\mathfrak{h}}, \underline{\mathfrak{h}}] \subseteq \underline{\mathfrak{h}}$, since any $\mathscr{Y} \in \underline{\mathfrak{h}}$ can be written as a linear combination of $\mathrm{L}(X_i)$ for $X_i \in \mathfrak{h}$. Thus, invoking Frobenius' theorem, take H to be the integral manifold of $\underline{\mathfrak{h}}$ through e,[12] i.e. choose a chart $f : G \to \mathbb{R}^d$ near e (where $d = \dim(G)$) sending $\underline{\mathfrak{h}} \to \langle \partial_1, \dots, \partial_k \rangle$, and let H be the group generated by $f^{-1}(x_1, \dots, x_k, 0, \dots, 0)$.

In general, H is dense in G, but we can topologize it differently: define a system of neighborhoods for $e \in H$ as $\{f^{-1}(V) \mid V$ is an open neighborhood of $f(e) \in \mathbb{R}^2\}$. Thus, H becomes a topological group, and using f, a Lie group, and the inclusion is smooth. There are several things to check here, but the point is that Frobenius' theorem does the hard work. ⊠

Now, given some isomorphism $\varphi : \mathfrak{g} \to \mathfrak{g}'$ of Lie algebras, let $\mathfrak{h} = \mathrm{graph}(\varphi) \subseteq \mathfrak{g} \times \mathfrak{g}'$, so by the above one has $H \to G \times G'$ with Lie algebra $\mathfrak{h}$ and such that the projections $H \to G$ and $H \to G'$ are diffeomorphisms near e. Thus, H locally gives a graph of a diffeomorphism $G \to G'$, and since H is a subgroup of $G \times G'$, then this diffeomorphism also locally preserves multiplication and inversion. This implies Lie's theorem.

We're not really going to use the above result, but there are two very useful facts coming from this viewpoint.

Theorem 7.3. *Suppose G is a Lie group and $H \subseteq G$ is a closed subgroup. Then, H is a submanifold of G, and thus a Lie subgroup.*

Corollary 7.4. *If $G \xrightarrow{\varphi} G'$ is a continuous group homomorphism of Lie groups, then it's smooth.*

This last corollary is particularly useful when discussing continuous representations of compact Lie groups.

Proof of Theorem 7.3. This proof will find an $\mathfrak{h} \subset \mathfrak{g}$ such that $H = \exp(\mathfrak{h})$ near e; that is, for some neighborhood V of 0 in $\mathfrak{g}$, $\exp(\mathfrak{h} \cap V) = H \cap \exp(V)$. This is equivalent to exp near e identifying $\mathfrak{h}$ with a subspace of $\mathfrak{g}$ (i.e. this is a coordinate chart in which $\mathfrak{h}$ is flat). This implies that H is a submanifold, and then checking the group operations are smooth is easy.

To produce $\mathfrak{h}$, look at how H approaches e. Let $\log : G \to \mathfrak{g}$ be a local inverse to the exponential, and let $\mathfrak{h}$ be the set of all limits in $\mathfrak{g}$ of sequences of the form $t_n \log(h_n)$ with $t_n \in \mathbb{R}$ and $h_n \to e$ (intuitively, the t_n allow one to rescale the sequence if necessary).

Then, if $X \in \mathfrak{h}$, then $\exp(X) \in H$, because we can write $X = \lim_{n\to\infty} t_n \log(h_n)$, but $h_n \to e$, so $t_n \to \infty$ (unless $X = 0$, but this case is trivial). Let m_n be the nearest integer to t_n, so that

$$\exp(X) = \lim_{n\to\infty} (\exp(t_n \log(h_n))) = \lim_{n\to\infty} \underbrace{h_n^{m_n}}_{\text{in } H} (\exp(t_n - m_n) \log h_n).$$

Since $t_n - m_n \leq 1$ for all n and H is closed, this implies $\exp(X) \in H$.

Next, we have to check that $\mathfrak{h}$ is a subspace of $\mathfrak{g}$; it's closed under scalars by definition, and if $X, Y \in \mathfrak{h}$, then $tX, tY \in \mathfrak{h}$ for all $t \in \mathbb{R}$, so $\exp(tX)\exp(tY) \in H$. But $\log(\exp(tX)\exp(tY)) = X + Y + \cdots$ (higher-order terms), so $X + Y \in H$ after taking a suitable sequence and rescaling.

Finally, we will need to show next lecture that $H = \exp(\mathfrak{h})$ locally, so to speak. This begins by picking a $w \subseteq \mathfrak{g}$ transverse to $\mathfrak{h}$, and so on…

8. Covering Groups: 4/16/14

Last time, we were in the process of showing that if G is a Lie group and $H \subseteq G$ is a closed subgroup, then H is a submanifold of G and thus a Lie subgroup. We defined $\mathfrak{h}$ to be the set of limits of sequences of the form $t_n \log(h_n)$, where $t_n \in \mathbb{R}$ and $h_n \to e$ in G, a subspace of $\mathfrak{g}$. If $X \in \mathfrak{h}$, then $\exp(X) \in H$.

[12] The theorem technically only gives a local result, but it can and should be extended globally.

Then, we can show that there's a neighborhood V of 0 in $\mathfrak{h}$ such that $\exp(V)$ is a neighborhood of the identity in H; then, by translation, one gets a similar chart around any $x \in H$, so H is in fact a submanifold.

Pick a transversal subspace $w \subseteq \mathfrak{g}$, so that $w \oplus \mathfrak{h} = \mathfrak{g}$. Then, the map $(Y \in w, X \in \mathfrak{h}) \mapsto \exp(X)\exp(Y)$ has derivative $(Y, X) \mapsto Y + X$; since w and $\mathfrak{h}$ are transversal, then this is an isomorphism $w \oplus h \xrightarrow{\sim} \mathfrak{g}$, so $(Y, X) \mapsto \exp(X)\exp(Y)$ is a local diffeomorphism.

Suppose there exists a sequence $h_m \in H$ such that $h_m \to e$, but such that $\log(h_m) \notin H$. Then, we can write $h_m = \exp(Y_m)\exp(X_m)$ with $Y_m \neq 0$ (since $\log(h_m) \notin H$), so $\exp(Y_m) = h_m \exp(-X_m) \in H$ (since $X_m \in \mathfrak{h}$), and thus any rescaled limit of the Y_m, i.e. $\lim t_m y_m$ with $t_m \in \mathbb{R}$, exists. (This has to do with the compactness of w.) Thus, it belongs to w as well as $\mathfrak{h}$, so it must be zero.

This makes it much easier to tell when something is a Lie subgroup (e.g. $\mathrm{O}(n)$, $\mathrm{U}(n)$).

Note that in Corollary 7.4, continuity is necessary; for example, if $\varphi \in \mathrm{Aut}(\mathbb{C})$, then it induces discontinuous group automorphisms on $\mathrm{GL}_n(\mathbb{C})$, and so on.

Proof of Corollary 7.4. $\mathrm{graph}(\varphi) = \{(x, \varphi(x))\} \subseteq G \times G'$ is a closed subgroup (which is a topological argument), so it's a subamanifold. Thus, the following diagram commutes.

$$\begin{tikzcd}
\mathrm{graph}(\varphi) \arrow[r, "\mathrm{proj}_2"] \arrow[d, "\mathrm{proj}_1"'] & G' \\
G \arrow[ur, "\varphi"] &
\end{tikzcd}$$

One can check that proj_1 is a diffeomorphism (by verifying that it has full rank on tangent spaces), so $\varphi = \mathrm{proj}_2 \circ \mathrm{proj}_1^{-1}$ is smooth. ⊠

Covering Groups. By the end of this lecture, we should be able to construct a Lie group that isn't (isomorphic to) a closed subgroup of a matrix group.

Definition 8.1. A covering map $\pi : X \to X'$ of topological spaces is a surjective local homeomorphism with discrete fibers at every point (i.e. the preimage $\pi^{-1}(x')$ for any $x' \in X'$, there's a neighborhood U of x such that $\pi^{-1}(U)$ is the disjoint union of open subsets of X, each of which maps homeomorphically onto U).

The standard example is $\mathbb{R} \to S^1$ given by $x \mapsto e^{2\pi i x}$.

Proposition 8.2. *Suppose G and G' are connected Lie groups. If $\pi : G' \to G$ is a covering map of topological spaces, then G' is a Lie group in a unique way such that π is a homomorphism of Lie groups.*

Example 8.3. $\pi_1(\mathrm{SL}_2(\mathbb{R}), e) \cong \mathbb{Z}$, so let $\widetilde{\mathrm{SL}_2(\mathbb{R})}$ be its universal covering; this is a Lie group that we will show is not isomorphic to a closed subgroup of $\mathrm{GL}_N(\mathbb{C})$. This has nothing to do with the fundamental group being infinite, even; it's still true for $\mathrm{SL}_n(\mathbb{R})$ for $n > 2$, yet in this case $\pi_1(\mathrm{SL}_n(\mathbb{R}), e) \cong \mathbb{Z}/2$ (so it's a double covering). ◀

This kind of behavior doesn't happen for compact groups; to be precise, the universal covering of a compact Lie group is still a matrix group, e.g. $\mathrm{SO}_n(\mathbb{R})$ is double covered by the spin group.

Proof of Proposition 8.2. First, we will establish the convention that when computing the fundamental group, the basepoint will always be the identity, i.e. $\pi_1(G) = \pi_1(G, e)$.

Claim. $\pi_1(G)$ is abelian.

Proof. The multiplication map $G \times G \to G$ induces $\pi_1(G \times G) \to \pi_1(G)$, but $\pi_1(G \times G) = \pi_1(G) \times \pi_1(G)$, so we have a map $f : \pi_1(G) \times \pi_1(G) \to \pi_1(G)$. Since multiplication is the identity when one coordinate is e (that is, $e \cdot g \mapsto g$ and such), then $f(x, e) = x$ and $f(e, y) = y$, so $f(x, y) = xy$. Thus, $(x, y) \mapsto xy$ is a homomorphism $\pi_1(G) \times \pi_1(G) \to \pi_1(G)$, which means that $\pi_1(G)$ must be abelian. ⊠

Now, let $f : G' \to G$ be a covering, and fix some $e' \in G'$ such that $f(e') = e$. Then, there's a unique way to lift the group operations from G to G'; the proof will demonstrate multiplication, and then inversion is the same. Basically, we want to complete this diagram by filling in the yellow arrow:

$$\begin{tikzcd}
G' \times G' \arrow[r, dashed] \arrow[d, "(f,f)"] & G' \arrow[d, "f"] \\
G \times G \arrow[r, "\text{mult.}"] & G
\end{tikzcd}$$

There's a simple criterion for when one can lift maps to a covering: suppose $f:\widetilde{Y}\to Y$ is a covering map and $\alpha : X\to Y$ is continuous. Then, α lifts to an $\widetilde{\alpha}$ such that the following diagram commutes precisely when $\alpha_*\pi_1(X)\subseteq f_*\pi_1(Y)$. This follows from the definition of a covering.

$$\begin{tikzcd} & \widetilde{Y} \arrow[d, "f"] \\ X \arrow[ur, dashed, "\widetilde{\alpha}"] \arrow[r, "\alpha"] & Y \end{tikzcd}$$

However, everything here is done in the category of pointed, connected topological spaces, i.e. we assume f and α preserve basepoints and all of the relevant spaces are connected.

Thus, back in Lie-group-land, we want to lift $\xi = \mathrm{mult}\circ(f,f): G'\times G'\to G$ to some map $\xi : G'\times G'\to G'$. In our course, all basepoints are the identity: e for G, and the specified e' for G'. Then,

$$\begin{aligned} \xi_*\pi_1(G'\times G') &= [\mathrm{mult}]_*(f_*\pi_1(G')\times f_*\pi_1(G')) \\ &= f_*\pi_1(G') + f_*\pi_1(G') \\ &\subseteq f_*\pi_1(G'). \end{aligned}$$

Thus, ξ lifts to ξ, so there's a smooth lift, and now one needs to check that it satisfies the group laws. This is a bunch of chasing axioms, but for example, for associativity, $(xy)z$ and $x(yz)$ agree after projection to G, so $x(yz) = ((xy)z)a(x,y,z)$ for some continuous a that is in $\ker(f)$. But since $\ker(f)$ is discrete, then $a = e'$ (since everything must preserve basepoints, so e must become the identity). ⊠

In particular, if G is a connected Lie group, then its universal covering $\widetilde{G}$ also has the structure of a Lie group, and is simply connected.

Proposition 8.4. *If G and H are Lie groups and G is simply connected,*[13] *then any homomorphism of Lie algebra $\varphi : \mathfrak{g}\to\mathfrak{h}$ is the derivative of some smooth homomorphism of Lie groups $\Phi : G\to H$.*

This is a pretty powerful result: we already knew that Φ existed locally, but if G is simply connected, then we're allowed to extend it globally. (This can in general be done with the universal cover of G, even when G isn't simply connected.)

As a degenerate example, if $G = H = \mathbb{R}/\mathbb{Z}$ (which is not simply connected), then $x\mapsto x\sqrt{2}$ on their respective Lie algebras doesn't lift to a map $G\to H$. However, this will always lift to a map from the universal cover of G, and we do indeed have an induced map $\mathbb{R}\to H$.

Proof of Proposition 8.4. By the BCH formula, the rule $\exp(X)\overset{F}{\mapsto}\exp(\varphi(X))$ is at least a local homomorphism, i.e. it goes from an open neighborhood U of $e\in G$ to an open neighborhood V of $e\in H$ and such that for all $u,u'\in U$, $F(u)F(u') = F(uu')$ whenever both sides make sense (e.g. $uu'\notin U$, or $F(u)F(u')\notin V$ would be examples of not making sense).

The way to turn this into a global homomorphism is the same way one does analytic continuation of a function. Let $\gamma : [0,1]\to G$ be a path such that $\gamma(0) = e$. Define

$$\Phi(\gamma) = F(\gamma(0)^{-1}\gamma(\varepsilon))F(\gamma(\varepsilon)^{-1}\gamma(2\varepsilon))\cdots F(\gamma(1-\varepsilon)^{-1}\gamma(1)),$$

where ε is small. That F is a local homomorphism means that $\Phi(\gamma)$ is independent of ε when it's sufficiently small, and in fact $\Phi(\gamma)$ depends only on the endpoint $\gamma(1)$, so the homomorphism we want is $\gamma(1)\mapsto\Phi(\gamma)$. Why's this? If one has two paths γ and γ' with the same endpoints, there's a homotopy γ_t between them, since G is simply connected. Thus, $\Phi(\gamma_t)$ is locally constant in t, so $\Phi(\gamma) = \Phi(\gamma_t) = \Phi(\gamma')$ (which is the same idea: break it into very small subsections). ⊠

Now, we can answer why $\widetilde{\mathrm{SL}_2(\mathbb{R})}$ isn't isomorphic to a closed subgroup of a matrix group. This is one of the many applications of Weyl's unitary trick: suppose there is a continuous (and therefore smooth) homomorphism $f:\widetilde{\mathrm{SL}_2(\mathbb{R})}\to\mathrm{GL}_n(\mathbb{C})$; then, one has a map $\mathrm{d}f : \mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}))\to\mathrm{Lie}(\mathrm{GL}_N(\mathbb{C}))$, which is real linear. But then, one can complexify it: $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}))\otimes_{\mathbb{R}}\mathbb{C} = \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$, given by the inclusion $\mathrm{SL}_2(\mathbb{R})\to\mathrm{SL}_2(\mathbb{C})$ (the exact details of which will be on the homework). Thus, one can extend $\mathrm{d}f$ to a complex linear map $\mathrm{d}f_{\mathbb{C}} : \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))\to\mathrm{Lie}(\mathrm{GL}_N(\mathbb{C}))$ that

[13]This means that $\pi_1(G)$ is trivial, and in particular that it is connected.

is a homomorphism of Lie algebras. But $\mathrm{SL}_2(\mathbb{C})$ is simply connected, so there's a map $f_{\mathbb{C}} : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$ with derivative extending that of f, which will imply that f must factor through $f_{\mathbb{C}}$, and in particular through $\mathrm{SL}_2(\mathbb{R})$, which will have interesting consequences next lecture.

The clever trick is that $\mathrm{SL}_2(\mathbb{C})$ is simply connected, but $\mathrm{SL}_2(\mathbb{R})$ isn't.

9. Haar Measure: 4/18/14

"What does 'useful' mean? There are people who use this… somewhere."

Last time, we were in the middle of showing that the universal covering $G \xrightarrow{\pi} \mathrm{SL}_n(\mathbb{R})$ (with $n \geq 2$) is a Lie group G that doesn't embed as a closed subgroup of $\mathrm{GL}_N(\mathbb{C})$ for any N. We have already that

$$\pi_1(\mathrm{SL}_n(\mathbb{R})) = \begin{cases} \mathbb{Z}, & n = 2 \\ \mathbb{Z}/2, & n > 2. \end{cases}$$

Continuation of the proof. The proof continues: if $\varphi : G \to \mathrm{GL}_N(\mathbb{C})$ is continuous (which will imply that it's C^∞), then we'll show that φ factors through $\mathrm{SL}_n(\mathbb{R})$, and therefore cannot be injective. This is because $\mathrm{d}\varphi : \mathrm{Lie}(G) \to \mathrm{Lie}(\mathrm{GL}_n(\mathbb{C})) \cong M_n(\mathbb{C})$, but $\mathrm{Lie}(G) \cong \mathrm{Lie}(\mathrm{SL}_n(\mathbb{R}))$ (since π is a local homomorphism, and the Lie algebra is a local construction). Thus, the map can be extended complex-linearly to a homomorphism of Lie algebras $\mathrm{d}\varphi_{\mathbb{C}} : \mathrm{Lie}(\mathrm{SL}_n(\mathbb{C})) \to \mathrm{Lie}(\mathrm{GL}_N(\mathbb{C}))$.[14]

Since $\mathrm{SL}_n(\mathbb{C})$ is simply connected, then $\mathrm{d}\varphi_{\mathbb{C}}$ is the derivative of a $\Phi : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$ (which we saw last time). Thus, the composition $G \xrightarrow{\pi} \mathrm{SL}_n(\mathbb{R}) \xrightarrow{\Phi} \mathrm{GL}_N(\mathbb{C})$ has the same derivative as $\varphi : G \to \mathrm{GL}_N(\mathbb{C})$, so they must be equal. This fact in question is that if $\varphi_1, \varphi_2 : G \to G'$ are homomorphisms of Lie groups with the same derivative and if G_1 is connected, then they're equal, which is true because they're equal near e thanks to the exponential map, and an open neighborhood of e generates an open-and-closed subgroup, which thus must be G. Thus, φ factors through $\mathrm{SL}_n(\mathbb{R})$. ⊠

Another way to word this is that no finite-dimensional representation can tell the difference between G and $\mathrm{SL}_n(\mathbb{R})$, much like how $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{R})$, and SU_n all have the same representation theory. This, like the above, uses Weyl's unitary trick.

Let's compute the fundamental group of $\mathrm{SL}_2(\mathbb{R})$. If a matrix in $\mathrm{SL}_2(\mathbb{R})$ is written $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then once one chooses $(c, d) \in \mathbb{R}^2 \setminus \{0\}$, then a and b are determined up to a line, so this space is contractible. Thus, $\mathrm{SL}_2(\mathbb{R}) \sim \mathbb{R}^2 \setminus \{0\} \sim S^1$ (where $\sim$ denotes homotopy equivalence), so the fundamental group is $\mathbb{Z}$. Similarly, $\mathrm{SL}_2(\mathbb{C})$ is simply connected because it's homotopy equivalent to $\mathbb{C}^2 \setminus \{0\} \sim S^3$, which is simply connected.

In more generality, one uses *fibrations* for larger n. For example, SO_n acts on S^{n-1}, and the stabilizer of any point is isomorphic to SO_{n-1}; thus, there's a fibration $\mathrm{SO}_{n-1} \to \mathrm{SO}_n \to S^{n-1}$, leading to a long exact sequence of homotopy groups π_j. This allows one to easily compute the fundamental groups of all Lie groups that are matrix groups, and for reference, they're listed here: $\pi_1(\mathrm{SO}_n) = \mathbb{Z}/2$ when $n > 2$ and $\pi_1(\mathrm{SO}_2) = \mathbb{Z}$; $\pi_1(\mathrm{SU}_n) = \pi_1(\mathrm{SL}_n(\mathbb{C}))$ is the trivial group for $n \geq 2$, and $\pi_1(U_n) = \pi_1(\mathrm{GL}_n(\mathbb{C})) = \mathbb{Z}$.

Theorem 9.1. *Let G be a Lie group; then, up to scaling, there is a unique left-invariant measure on G (i.e. for all $S \subseteq G$ and $g \in G$, $\mu(gS) = \mu(S)$). The same is true for right-invariance.*

By "measure" this theorem means a Radon measure, on σ-algebras and Borel sets and so on. But it's more useful to think about this as giving an integration operator from $C_C(G) \to \mathbb{R}$ (i.e. from the set of compactly supported functions on G); that is, a continuous functional such that

$$\int_G \mathsf{L}_h f \, \mathrm{d}g = \int_G f \, \mathrm{d}g,$$

where $f \in C_C(G)$ and $\mathsf{L}_h f$ denotes a left-translate by an $h \in G$.

For a simple example, when $G = (\mathbb{R}, +)$, the Haar measure is the Lesbegue measure, and the left-invariance tells us that this is invariant under translation.

In $\mathrm{GL}_n(\mathbb{R})$, one could intuitively use the induced Lesbegue measure μ_L from $M_n(\mathbb{R})$, but $\mu_L(gS) = \det(g)^n \mu_L(S)$, since this is in $\mathbb{R}^{n \times n}$, so there are n copies of the determinant; the point is it isn't invariant. If one instead takes

[14] This is also complex linear, but that's not crucial to the point of the proof.

$\mu_L/\det^n$, this becomes a left- and right-invariant measure. In general, the left- and right-invariant measures don't have to be the same, such as for the Haar measure on

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \in \mathbb{R} \right\},$$

but they do coincide for compact groups (as we will be able to show). But one can do better: if you normalize the measure, so that $\mu(G) = 1$, then there is a unique left-invariant *probability* measure on G, and it is also right-invariant. For the rest of the course, this probabilistic measure is what is meant by Haar measure unless otherwise specified.

The above is actually true for any locally compact topological space, but the proof is harder, and in any class of groups one might apply this to, one can just compute the Haar measure anyways.

Proof of Theorem 9.1. This proof looks constructive, but don't let it fool you: the "formula" is totally unhelpful.

To construct the desired measure μ, let $n = \dim(G)$ and $\mu = |\omega|$, where ω is a differential form of top degree (so that it can be integrated). Such an ω exists because it can be defined in terms of invariants: for some $\omega_0 \in \wedge^n (T_e G)^*$, let $\omega(g) = (l_{g^{-1}})^* \omega_0$. This is valid for the same reason left-invariant vector fields can be constructed from vectors in $T_e G$ when constructing the Lie algebra: once you're somewhere, you can be everywhere. Thus, ω is left-invariant, i.e. $(l_x)^* \omega = \omega$, so $\mu = |\omega|$ is left-invariant too.

Suppose ν is another left-invariant measure; then, it's absolutely continuous for μ, i.e. if S has zero measure for ν, then it does for μ. This fact is true only because μ is G-invariant. Thus, $\nu = f\mu$ for some $f \in L^2(\mu)$ by the Radon-Nicholson theorem, but this means f is left-invariant, so it must be constant. ⊠

If you really had to compute this, you could: let's make this explicit in a more complicated case than just the general linear group,

$$\mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{bmatrix} x & y \\ z & w \end{bmatrix} \mid xw - yz = 1 \right\}.$$

Then, the map $\mathrm{SL}_2(\mathbb{R}) \to \mathbb{R}^3$ sending $\left[\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right] \mapsto (x, y, z)$ is bijective away from $x = 0$. Then, the Haar measure in $\mathrm{SL}_2(\mathbb{R})$ will be a function of the Lesbegue measure on $\mathbb{R}^3$.

First, let's write down a left-invariant 3-form on $\mathrm{SL}_2(\mathbb{R})$. Though it's possible to do it systematically, in general it's better to guess and hope for the best. Let

$$g^{-1}\,\mathrm{d}g = \begin{bmatrix} w & -y \\ -z & x \end{bmatrix} \begin{bmatrix} \mathrm{d}x & \mathrm{d}y \\ \mathrm{d}z & \mathrm{d}w \end{bmatrix} = \begin{bmatrix} w\,\mathrm{d}x - y\,\mathrm{d}z & w\,\mathrm{d}y - y\,\mathrm{d}w \\ -z\,\mathrm{d}x + x\,\mathrm{d}z & -z\,\mathrm{d}y + x\,\mathrm{d}w \end{bmatrix}.$$

This is left-invariant, because when left-multiplying by h, it becomes $g^{-1}h^{-1}h\,\mathrm{d}g$, which cancels out nicely. Thus, the entries give four left-invariant one-forms (though they're not linearly independent), three of which span the space of one-forms. Thus, one can make an invariant 3-form by wedging any three of them together, and the specific choice only matters up to a constant factor. For example, because $\mathrm{d}(xw - yz) = 0$, then it's possible to simplify

$$(w\,\mathrm{d}y - y\,\mathrm{d}z) \wedge (w\,\mathrm{d}y - y\,\mathrm{d}u) \wedge (-z\,\mathrm{d}x + x\,\mathrm{d}z)$$

into a big mess of wedges which eventually becomes

$$\frac{\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z}{x}.$$

In other words, away from $x = 0$, the Haar measure looks like $1/x$ times the Lesbegue measure.

This is systematic in that it works for any matrix group, but it's worth thinking about why it's so simple in the end, and also worth investigating for $\mathrm{SL}_n(\mathbb{R})$.

In the case of compact topological groups, there's a more abstract proof illustrating another way to think about the Haar measue. A probability measure is, after all, a notion of a random element of G, and $\mu(S)$ is the probability that a random element lies in the set S. So how might we actually draw random elements from this distribution? The goal is to produce a sequence $g_1, g_2, \dots$ in G such that, for nice $S \subseteq G$, $\#\{1 \le i \le n : g_i \in S\}/n \to \mu(S)$. For example, on

$$\mathrm{SU}_2 = \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}, |a|^2 + |b|^2 = 1 \text{ for } a, b \in \mathbb{C} \right\},$$

which can be embedded as $S^3 \subseteq \mathbb{R}^4$, the "area measure" on this sphere ends up being the Haar measure (though this is three-dimensional area).

Proposition 9.2. *Choose some $x_1, \dots, x_n \in G$ that generate a dense subgroup (i.e. $\overline{\langle x_1, \dots, x_n \rangle} = G$).*[15] *Then, a random sequence in G can be obtained by $g_{i+1} = g_i x_r$, where $r = 1, \dots, n$ is chosen randomly, and this sequence samples from the Haar measure.*

10. Applications to Representation Theory: 4/21/14

Throughout today's lecture, let G be a compact group; we'll be applying these results specifically to compact Lie groups, but the proofs are the same in this greater generality.

Last time, we constructed the Haar measure with differential forms. Now, we can give a different construction for compact groups: choose $g_1, \dots, g_N \in G$ that span a dense subset of G.[16] Then, pick some large K and for $1 \le k \le K$, setting $g = x_1, \dots, x_k$, where each x_i is chosen uniformly at random from $\{g_1, \dots, g_N\}$, samples at random from the Haar measure.

In order to formulate this more precisely, given this set $\{g_1, \dots, g_N\}$, define operators L and R for left averaging with it:

$$Lf(g) = \frac{1}{N}(f(g_1 g) + \cdots + f(g_N g))$$
$$Rf(g) = \frac{1}{N}(f(g g_1) + \cdots + f(g g_N))$$

Proposition 10.1. *If $C(G)$ denotes the set of continuous functions on G (which has a topology induced from the supremum norm), then for any $f \in C(G)$,*

$$\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} L^k f(e) = \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} R^k f(e) = \int f \,\mathrm{d}\mu,$$

where μ is the Haar measure.

This will be shown to define a left- and right-invariant measure, and then that such a measure is unique. Specifically, we'll show that

$$\frac{1}{K}\sum_{k=1}^{K} L^k f(e) \longrightarrow c,$$

where c is some constant function (and the convergence is in the topology induced by the supremum norm), and the same for $R^k f$.

Proof of Proposition 10.1. Let $f_K = \sum_1^K L^k f$ and $f'_K = \sum_1^K R^k f$.

Note that if $F \in C(G)$ is fixed by left averaging, then F must be constant,[17] because it's a function on a compact space, so look at the x such that $|F(x)|$ is maximized; then, $F(g_i x) = F(x)$ for all i, and then iterating, $F(g_i g_j x) = F(x)$, and so on. But since $\overline{\langle g_1, \dots, g_N \rangle} = G$, then F must be constant.

The collection $\{f_K\}$ is precompact, which is to say that it has compact closure in $C(G)$, and (more useful for this proof) it has a convergent subsequence. Then, the proof boils down to checking that they're equicontinuous, i.e. continuous in a uniform manner in K. This uniformity and the existence of a convergent subsequence means that the overall limit must exist.

But this ends up being true by a translation argument, and the argument for f'_K is identical, so there's a subsequence of $\{f_K\}$ converging to a limit: $f_{K_n} \to f_\infty$. This f_∞ must be a constant, because

$$Lf_K - f_K = \frac{L^{k+1} f - f}{K} \le \frac{2\|f\|_\infty}{K} \to 0,$$

and thus $Lf_\infty = f_\infty$, so f_∞ is a constant, as seen above. Thus, any convergent subsequence of $\{f_K\}$ converges to a constant function on G — though the constant can depend on f.

Now, let's look at any convergent subsequence of right averages: $f'_{K_n} \to f'_\infty$, which by the same reasoning must be constant. But left and right averaging commute, so take this f'_∞ and left-average it, so it must be equal to f_∞.

[15] This is allowed to be an infinite sequence of x_i, if you want.

[16] Note that this actually can't be done in the case of an arbitrary compact topological group, though it is possible for compact Lie groups. This isn't too much of a setback, however; the proof can easily be adapted to an infinite spanning set by choosing an arbitrary G-valued probability measure and replacing the averaging operators with integrals across this measure.

[17] This section of the proof resembles the proof of the maximum modulus principle.

Thus, every convergent subsequence of f'_K converges to f_∞, so since $\{f'_K\}$ is precompact, then $f'_K \to f_\infty$ as well. Then reversing left and right, $f_K \to f_\infty$ too.

Let $\nu(f)$ be the constant induced by a given $f \in C(G)$, so that $f \mapsto \nu(f)$ is a continuous linear functional on $C(G)$, so it's a measure, and since $\nu(f) \geq 0$ for $f \geq 0$ and $\nu(1) = 1$, then ν is a probability measure. Furthermore, by the construction above, ν is both left- and right-invariant.

If μ is any left-invariant probability measure, then apply it to $(1/K)\sum_1^K L^k f$, and once all of the left-averaging is done, $\mu(f) = \nu(f)$. ⊠

This argument is much harder in the noncompact case, and considerably less useful.

Basic Results from Representation Theory. Moving into representation theory, for the rest of the course, G will generally denote a compact group, and usually a Lie group. Right now, it doesn't need to be Lie, at least until we start differentiating representations.

Definition 10.2. A representation of a topological group G is a continuous homomorphism $\rho : G \to \mathrm{GL}_N(\mathbb{C})$.

If G is Lie, then ρ is automatically smooth.

All of the main results in the representation theory of finite groups still hold. These two will be particularly useful.

Proposition 10.3.

(1) *A representation is irreducible if it has no nontrivial subrepresentations, and every representation is a direct sum of irreducible subrepresentations.*
(2) *The characters $g \mapsto \mathrm{tr}(\rho(g))$ of irreducible representations form an orthonormal basis for the Hilbert space of class functions $\{f \in L^2(G,\mu) \mid f(xgx^{-1}) = f(g)\}$ (where μ is, as always, the Haar measure).*

These are both easy to prove, but with one concern: it's not clear that there are any representations, irreducible or not, besides the trivial one. With subgroups of matrix groups, it's possible to make it work, but it's a bit harder abstractly.

Proof of Proposition 10.3, part (1). Let V be a representation of G, i.e. a finite-dimensional vector space with a continuous action of G. Then, there exists a G-invariant inner product on V: choose any inner product $\langle v, v'\rangle$, and average it over G (which is the whole point of the Haar measure):

$$[v, v'] = \int_{g\in G} \langle gv, gv'\rangle \, \mathrm{d}g.$$

(Notice that this is identical to the proof in the case of finite groups, but uses an integral over G rather than a sum.) Then, $[hv, hv'] = [v, v']$ for $h \in G$.

Thus, if $W \subseteq V$ is a subrepresentation of V, then $W^\perp$ is also, where $\perp$ is with respect to $[v, v']$. Thus, $V = W \oplus W^\perp$, and continuing in this way, V is a direct sum of irreducibles. ⊠

We also have Schur's lemma, with exactly the same proof.

Lemma 10.4 (Schur). *If V and W are irreducible representations of G, then any G-homomorphism (i.e. a linear map commuting with the action of G) $T : V \to W$ is 0 if $V \ncong W$, or $\lambda \cdot \mathrm{Id}$, if $V \cong W$.*

The proof is short, and looks at the eigenspaces of T, or its image and kernel as subrepresentations (which don't have many options, since V and W are irreducible).

Proof of Proposition 10.3, part (2). We'll first show that $\langle \chi_V, \chi_W\rangle = 0$ when $V \ncong W$ and both are irreducible (where χ_V is the trace of the action of $g \in G$ on V, and χ_W is analogous). If one has G-invariant inner products on V and W, then for any $v_0 \in V$ and $w_0 \in W$, set $S(v) = \langle v, v_0\rangle w_0$, so that $S : V \to W$ is linear (and all linear maps can be obtained in this way). Averaging it across G produces

$$T(v) = \int_G gSg^{-1}(v)\, \mathrm{d}g$$

for $v \in V$, so $T : V \to W$ is G-invariant, as $T(hv) = hT(v)$ for any $h \in G$, and by Schur's lemma, $T = 0$.

This will imply orthogonality: $gSg^{-1}(v) = \langle g^{-1}v, v_0\rangle g w_0$, so what this is saying is that

$$\int_{g\in G} \langle g^{-1}v, v_0\rangle g w_0 \,\mathrm{d}g = 0.$$ [18]

Then, take the inner product with any $w \in W$: by the G-invariance of the inner product, $\langle g^{-1}x, y\rangle = \langle x, gy\rangle = \overline{\langle gy, x\rangle}$, so

$$\int_{g\in G} \overline{\langle g v_0, v\rangle}\langle g w_0, w\rangle \,\mathrm{d}g = 0.$$

But if v_i is an orthonormal basis for V, then

$$\chi_V(g) = \sum_i \langle g v_i, v_i\rangle,$$

and similarly for an orthonormal basis w_i of W, but we've just shown that these terms are pointwise orthogonal, so $\langle \chi_V, \chi_W\rangle = 0$. ⊠

This proof used compactness implicitly in the Haar measure, and to get that gSg^{-1} converges to a continuous linear functional. Thus, it is sometimes generalizable.

11. The Peter-Weyl Theorem: 4/23/14

Last time we showed that if G is a compact group, then every representation is a finite sum of irreducibles, and the characters of irreducible representations form an orthonormal basis for the space of class functions in $L^2(G)$ (though we still need to show that if V is irreducible, then $\langle \chi_V, \chi_V\rangle = 1$).

These together imply that a representation V of G is determined up to isomorphism by its character, and that the multiplicity of an irreducible representation W in V is $\langle \chi_V, \chi_W\rangle$. Orthogonality comes from the fact that if V and W are irreducible and $V \not\cong W$, then $\mathrm{Hom}_G(V, W) = 0$, and we saw that $\langle \chi_V, \chi_W\rangle = \dim \mathrm{Hom}_G(V, W)$ in general.

What we still haven't shown, though, is that the characters span a dense subset of $L^2(G)$, so that there exist representations other than the trivial one; this will be one of the consequences of the Peter-Weyl theorem. This is important; there exist groups where elements are only conjugate to powers of themselves, and thus there are no finite-dimensional representations.

If $n \geq 3$, then $\pi_1(\mathrm{SO}_n) \cong \mathbb{Z}/2$, so the universal cover G of SO_n is a double cover. This is a compact Lie group, and it's not immediately clear how to produce representations of G that don't come from SO_n. In the case of finite groups, it's possible to look at the space of functions acting on themselves, but here it's necessary to have some sort of finite-dimensional point of view.

Definition 11.1. If f is a continuous function on a compact group G, then f is called left-finite if its left translates ($g \mapsto f(xg)$ for some $x \in G$) span a finite-dimensional vector space; right-finite functions are defined in the analogous way.

Theorem 11.2 (Peter-Weyl)**.**

(1) *The following are equivalent for a compact group G and a function f on G.*
   - *f is left-finite.*
   - *f is right-finite.*
   - *$f(g) = \langle gu, v\rangle_V$ for a finite-dimensional G-representation V and $u, v \in V$.*

(2) *Finite functions are dense in $L^2(G)$.*

Sometimes, the term "the Peter-Weyl theorem" is only used for part 2 of this theorem.

The point of this theorem is that there are "enough" finite-dimensional subspaces arising from representations, so there are actually nonzero representations.

Example 11.3. If $G = \mathrm{U}_n$, then f is finite iff it is polynomial in the entries g_{ij} and $\overline{g_{ij}}$ as a matrix, though this is not obvious. ◀

[18] This is an integral on a finite-dimensional vector space, so take it componentwise.

The Peter-Weyl theorem implies that the span of characters is dense in $L^2(G)$, so given any class function F and an $\varepsilon > 0$, there is a finite-dimensional representation V and $u, v \in V$ such that

$$||F - \langle gu, v\rangle||_{L^2} \leq \varepsilon.$$

F is already invariant under conjugation, so averaging under the conjugation action doesn't do anything. And since conjugation is an isometry on $L^2(G)$, then it doesn't change the L^2-distance, so

$$\left\| F - \int\limits_{h\in \mathrm{G}} \langle hgh^{-1}uv\rangle \,\mathrm{d}h \right\|_{L^2} \leq \varepsilon.$$

This integral is spanned by characters: first, one can assume V is irreducible (and if not, split it into its irreducible components), so that for $u, v \in V$,

$$\int\limits_{h\in G} hgh^{-1}u \,\mathrm{d}h = \frac{\chi_V(g)}{\dim V}u.$$

In fact, since $\int_{h\in G} hgh^{-1}$ was averaged over G, then it's an endomorphism of V that commutes with G. Thus, by Schur's lemma, it's $\lambda \cdot \mathrm{Id}$, and λ can be computed with traces, as

$$\mathrm{tr}\left(\int_{h\in G} hgh^{-1}\right) = \chi_V(g) = \lambda \dim(V).$$

Proof of Theorem 11.2. For (1), suppose f is left-finite. Then, let $L_g f$ denote the left translate of f by g, i.e. $h \mapsto f(gh)$, and define $R_g f$ similarly. If V is the span of the left translates of f, then $f(g) = L_g f(e)$. Since L_g is a functional on a finite-dimensional space, then there exists a $\delta \in V$ such that $f(g) = \langle L_g f, \delta\rangle$.[19] Then, the other directions are analogous.

For (2), this is really easy if G is a closed subgroup of a matrix group (i.e. $G \subseteq \mathrm{GL}_N(\mathbb{C})$). This is because polynomials in g_{ij} and $\overline{g_{ij}}$ are finite functions (and in fact, thee are all of the finite functions, but this is harder to show), and these functions are already dense in $L^2(G)$ (via the L^∞ norm) by the Stone-Weierstrass theorem.

In the general case, we need a way to break up $L^2(G)$ into finite-dimensional subspaces, which involves producing a character $L : L^2(G) \to L^2(G)$ with finite-dimensional eigenspaces, and that commutes with the right action of G; then, the eigenspaces will consist of finite functions. There are various examples for more specific cases, e.g. the Laplacian for compact Lie groups, but for the general case we'll use something that looks like the operator that yielded the Haar measure $Lf(x) = (1/N)\sum_{i=1}^n f(g_i x)$.

Let δ be a continuous function on G, and define

$$L_\delta f(x) = \int_G \delta(g) f(gx) \,\mathrm{d}g.$$

In some sense, this is a convolution operator.[20]

Now, we have an operator $L_\delta : L^2 \to L^2$. It would be nice if it were self-adjoint, too. In the finite case for the Haar measure, the adjoint L^* sends $g_i \mapsto g_i^{-1}$, and is otherwise fine; thus, if $\delta(x) = \overline{\delta(x^{-1})}$, then L_δ is self-adjoint, and since δ is continuous, then L_δ is compact. That is, for all ε there's an operator F of finite rank (i.e. finite-rank image) such that $||L_\delta - F(v)|| \leq \varepsilon ||v||$, using the operator norm.

By the spectral theorem for compact, self-adjoint operators, L^2 is a direct sum of finite-dimensional eigenspaces (well, for $\lambda \neq 0$), and the nonzero eigenspaces correspond to finite functions. In particular, any $F \in \mathrm{L}^2$ can be approximated by a finite function $L_\delta f$ arbitrarily well (in the sense of L^2); as the approximation gets better, δ approaches the Dirac delta function, supported on smaller and smaller neighborhoods of the identity, but with total integral 1. ⊠

Some of the trickier aspects of this operator stuff are discussed in the textbook.

[19] Be careful; there really should be a factor of g^{-1} in the precise calculations, but that's all right for the proof.

[20] Another common definition uses instead $\delta(g^{-1})f(g^{-1}x)\,\mathrm{d}g$, but it doesn't actually matter.

Classifying Representations: an Overview. The course will continue by classifying representations of compact Lie groups, and, eventually, the Lie groups themselves. We'll start with SU_2, which is probably the most important part of the course, and can be written as

$$\mathrm{SU}_2 = \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\},$$

which is topologically equivalent to $S^3 \subset \mathbb{R}^4$. We'll also talk about the closely related SO_3, since SU_2 is its double cover, a universal cover. Thus, there is a continuous homomorphism $\mathrm{SU}_2 \to \mathrm{SO}_3$ with kernel $\{\pm 1\}$.[21] The point is, most things about SO_3 will follow from SU_2, and are a bit easier to visualize.

One can identify SU_2 with the group of quaternions with norm 1 (the norm of $x+yi+zj+wk$ is $x^2+y^2+z^2+w^2$) by sending

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \longmapsto a + bj,$$

since $a, b \in \mathbb{C}$. This does end up being a homomorphism, and if $a = x + iy$ and $b = z + iw$, then it gets sent to $x + iy + zj + wk$, and the condition $|a|^2 + |b|^2 = 1$ is equivalent to the condition of norm 1.

Now, given a quaternion $q \in \mathrm{SU}_2$, there's a rotation in SO_3 given by $xi + yj + zk \mapsto q(xi + yj + zk)q^{-1}$, which preserves $x^2 + y^2 + z^2$, but shuffles the terms. Additionally, the adjoint map $\mathrm{Ad}(g) = D(gxg^{-1})$, $\mathrm{Ad} : \mathrm{SU}_2 \to \mathrm{GL}(\mathrm{Lie}(\mathrm{SU}_2))$. BY differentiating the unitary condition, $\mathrm{Lie}(\mathrm{SU}_2) \subseteq \mathrm{Lie}(\mathrm{GL}_2(\mathbb{C})) = M_2(\mathbb{C})$, and in fact, it is the set of $X \in M_2(\mathbb{C})$ such that $X + \overline{X}^{\mathrm{T}} = 0$ and $\mathrm{tr}(X) = 0$, i.e.

$$\mathrm{Lie}(\mathrm{SU}_2) = \left\{ \begin{pmatrix} ia & b + ic \\ -b + ic & -ia \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\} \subseteq M_2(\mathbb{C}).$$

This is a three-dimensional vector space. Then, $\mathrm{Ad}(g) : X \mapsto gXg^{-1}$ preserves the quadratic form $\mathrm{tr}(X^2) = -2(a^2 + b^2 + c^2)$. In particular, the matrix $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right) \in \mathrm{SU}_2$ is sent to the action of rotation by 2θ in SO_3.

12. Representations of SO_2 and SU_3: 4/25/14

> *"At some point, you notice things and become depressed at how long ago they were discovered... this is an extreme example."*

Last time, we defined a map $\pi : \mathrm{SU}_2 \to \mathrm{SO}_3$, which is a double cover with kernel $\{\mathrm{Id}, -\mathrm{Id}\}$ in terms of quaternions. It could also be thought of as the map $\mathrm{Ad} : \mathrm{SU}_2 \to \mathrm{Aut}(\mathrm{Lie}(\mathrm{SU}_2))$, or $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right) \mapsto \mathrm{rot}(2\theta)$.

We want to understand how to integrate things, especially class functions, within SU_2 and thus SO_3. SU_2 is intuitively the complex unit sphere (and thus three-dimensional): the Haar measure on it is the same as the area measure of $S^3 \subseteq \mathbb{R}^4$, suitably normalized, and it's invariant because the area measure of S^3 is invariant under rotation.

Any element of SU_2 is conjugate to an element of the type

$$r_\theta = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix},$$

where θ is determined up to sign, so the conjugacy classes are parameterized by the rotation angle θ, with $\theta \in [0, \pi]$. Thus, if F is a class function, with $F(\theta) = F(r_\theta)$, then

$$\int_{\mathrm{SU}_2} F = \frac{2}{\pi} \int_0^\pi F(\theta) \sin^2\theta \, \mathrm{d}\theta. \tag{12.1}$$

On SO_3, this means that if one chooses a random rotation of three-dimensional space, there's a similar formula (given by the induced measure from SU_2) that states that the probability that the rotation angle φ is proportional to $\sin^2(\varphi/2)$, so it's unlikely to be small.

This will admit many generalizations.

Here's a proof of (12.1). Let $\sim$ denotes conjugacy and suppose

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \sim \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix},$$

[21]This means that $\mathrm{SU}_2 = \mathrm{Spin}(3)$; the spin groups are the universal covers of SO_n.

then one can take the trace $a+\overline{a}=2\cos\theta$, so $\mathrm{Re}(a)=\cos\theta$. This projects S^3 onto the real line (specifically, $[-1,1]$).[22] Unlike the three-dimensional case, he induced measure from S^3 depends on z: it becomes $\sqrt{1-z^2}\,\mathrm{d}z$. Letting $z=\cos\theta$, this becomes $\sin^2\theta\,\mathrm{d}\theta$, and then it must be normalized to become a probability measure, leading to (12.1).

Irreducible Representations. Let V_n be the space of homogeneous polynomials on $\mathbb{C}^2$ of degree n (i.e. those polynomials with only top-degree terms). This is an $(n+1)$-dimensional representation of SU_2, where $g\cdot P(gx)=P(x)$, so $gP(x)=P(g^{-1}x)$ — it's always slightly confusing for groups to act on functions

Claim. All of the V_n are irreducible, and give all of the irreducible representations of SU_2.

Proof. This can be checked with character theory. Write $V_n=\langle x^n, x^{n-1}y,\dots,y^n\rangle$, so the action of r_θ is

$$\begin{aligned} r_\theta(x^n) &= e^{in\theta}x^n\\ r_\theta(x^{n-1}y) &= e^{i(n-2)\theta}x^{n-1}y \end{aligned}$$

and so on, i.e. r_θ acts by the matrix

$$\begin{pmatrix} e^{in\theta} &&&&\\ & e^{i(n-2)\theta} &&&\\ && e^{i(n-4)\theta} &&\\ &&& \ddots &\\ &&&& e^{-in\theta} \end{pmatrix}.$$

Thus, the character is

$$\begin{aligned} \chi_{V_n}(r_\theta) &= e^{in\theta}+e^{i(n-2)\theta}+\cdots+e^{-in\theta}\\ &= \frac{e^{i(n+1)\theta}-e^{-i(n+1)\theta}}{e^{i\theta}-e^{-i\theta}}, \end{aligned} \tag{12.2}$$

since it's a geometric series. Therefore

$$\begin{aligned} \left\langle \chi_{V_n},\chi_{V_n}\right\rangle &= \frac{2}{\pi}\int_0^\pi \left(\frac{e^{i(n+1)\theta}-e^{-i(n+1)\theta}}{e^{i\theta}-e^{-i\theta}}\right)^2 \sin^2\theta\,\mathrm{d}\theta\\ &= \frac{2}{\pi}\int_0^\pi \frac{1}{4}|e^{i(n+1)\theta}-e^{-i(n+1)\theta}|^2\,\mathrm{d}\theta\\ &= 1, \end{aligned}$$

so V_n is irreducible.[23]

To show that these complete the irreducible representations boils down to Fourier analysis on S^1: something orthogonal to all of these would also have to be orthogonal to all functions of the form $\chi_{V_n}(\theta)\sin^2\theta$, but this can be attacked with Fourier analysis. ⊠

Thus, there's a unique irreducible representation of SU_2 for each $n\geq 1$.

One could actually have checked that the irreducible characters had the form (12.2) by pure thought, without knowing the structure of V_n, as in the first few lectures with U_n. The idea is that if W is an irreducible representation, then

$$\chi_W(r_\theta)=\sum_{m\in\mathbb{Z}} a_m e^{im\theta}$$

[22] Note that for $S^2\subseteq\mathbb{R}^3$ projecting down onto the real line, via $(x,y,z)\mapsto z$, the induced measure on $\mathbb{R}$ is $\mathrm{d}z/2$; this fact, in different words, was known to Archimedes!

[23] Note that there's a formula for the dimension:

$$\int_G \langle gv_1,v_2\rangle\langle gv_1',v_2'\rangle\,\mathrm{d}g=\frac{\langle v_1,v_1'\rangle\langle v_2,v_2'\rangle}{\dim V},$$

but for character calculations, this isn't all that important.

for some $a_m \geq 0$, which comes from restricting W to a representation of the diagonal subgroup $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right)$. Then, $\langle \chi_W, \chi_W \rangle = 1$ is equivalent to the following: if one defines b_m by

$$\left(\sum_{m\in\mathbb{Z}} a_m e^{im\theta}\right)\left(e^{i\theta} - e^{-i\theta}\right) = \sum_{m\in\mathbb{Z}} b_m e^{im\theta},$$

then $\sum b_m^2 = 2$, and this expression is odd in θ (determined by switching θ and $-\theta$).[24] But $r_\theta \sim r_{-\theta}$, so this shouldn't change anything, so in particular one can only have the formula (12.2), and then check that they're irreducible, as we did.

What happens to SO_3**.** The irreducible representations of SO_3 correspond bijectively to those of SU_2 on which $-I$ acts trivially, which is shown by pushing representations back and forth across $\pi : \mathrm{SU}_2 \to \mathrm{SO}_3$, whose kernel is $\{\pm I\}$. $-I$ acts on V_n by $(-1)^n$, so the V_n for n even descend to SO_3, giving an $(n+1)$-dimensional representation with character

$$\chi_{V_n}(\varphi) = e^{i(-n/2)\varphi} + e^{i(-n/2+1)\varphi} + \cdots + e^{i(n/2)\varphi},$$

where φ is the rotation angle in SO_3; the factor of two is because $\pi(r_\theta)$ is a rotation by the angle 2θ. Thus, SO_3 has a unique irreducible representation of each odd dimension. This is cool, but it would certainly be nice to have a more direct visualization — this ends up being the theory of spherical harmonics, which will provide an explicit model for these irreducibles.

Let P_n be the space of homogeneous, $\mathbb{C}$-valued[25] polynomials on $\mathbb{R}^3$ of dimension n, so that $P_0 = \langle 1 \rangle$, $P_1 = \langle x, y, z \rangle$, $P_2 = \langle x^2, y^2, z^2, xy, yz, xz \rangle$, and so forth. The dimensions are $1, 3, 6, 10, \ldots$ thus, these aren't irreducible for $n \geq 2$, as all of the irreducibles were shown to have odd dimension. However, we can still extract the irreducibles.

Let $\mathrm{rot}_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi & \\ \sin\varphi & \cos\varphi & \\ & & 1 \end{pmatrix}$ denote rotation by the angle φ about the z-axis, so that

$$\chi_{P_1}(\mathrm{rot}_\varphi) = 2\cos\varphi + 1 = e^{i\varphi} + e^{-i\varphi} + 1.$$

Computing higher χ_{P_n} in this basis is unpleasant, so change coordinates to $e_1 + ie_2$, $e_1 - ie_2$, and e_3, so that rotation acts as the diagonal matrix given by $(e^{i\varphi}, e^{-\varphi}, 1)$. In this basis,

$$P_2 = \langle (x+iy)^2, (x-iy)^2, z^2, (x+iy)(x-iy), (x+iy)z, (x-iy)z \rangle,$$

so rot_φ acts diagonally, via the matrix

$$\begin{pmatrix} e^{2i\varphi} & & & & & \\ & e^{-2i\varphi} & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & e^{i\varphi} & \\ & & & & & e^{-i\varphi} \end{pmatrix}.$$

Thus, the character becomes

$$\chi_{P_2}(\mathrm{rot}_\varphi) = e^{2i\varphi} + e^{-2i\varphi} + 2 + e^{-i\varphi} + e^{i\varphi}.$$

This looks *almost* like an irreducible character, but with a 2 instead of a 1, so P_2 must be the direct sum of a five-dimensional irreducible representation and a one-dimensional irreducible representation. $x^2 + y^2 + z^2$ is invariant under rotation, so one way to think of the five-dimensional representation is as the quotient of P_2 by $(x^2 + y^2 + z^2)$. Alternatively, it's the kernel of the Laplacian $\Delta = \partial_{xx} + \partial_{yy} + \partial_{zz}$ (i.e. things that are rotation-invariant), given by things of the form

$$\{ax^2 + by^2 + cz^2 + \cdots \mid a + b + c = 0\}.$$

In general, after computing characters, $P_n = P_{n-2} \oplus H_n$, where H_n is a $(2n+1)$-dimensional irreducible representation, called the "harmonic polynomials," and there are corresponding maps $P_{n-2} \to P_n$ given by multiplication by $(x^2 + y^2 + z^2)$ and $P_n \to P_{n-2}$ given by taking the Laplacian. However, these maps aren't inverses.

[24]This argument might look a little cloudy, but we'll see it again in a more general case later on in the course, where it will be given a more proper proof.

[25]Everything in the study of complex representations is also complex: vector spaces are over $\mathbb{C}$, spans are $\mathbb{C}$-valued, and so on.

This creates two realizations of the $(2n+1)$-dimensional irreducible representation of SO_3: as P_n/P_{n-2} (which isn't used very often) or as $\ker(\Delta)$ on P_n.[26]

One interesting consequence is that, as Hilbert spaces,

$$L^2(S^2) = \widehat{\bigoplus_n} H_n|_{S^2},$$

and $H_n|_{S^2}$ is the eigenspace for Δ_{S^2}.[27] $H_n|_{S^2}$ is known as the space of spherical harmonics, and there's plenty of classical literature about it.

It's also possible to start with S^2 and use this to develop the irreducible representations, but this means one has to think about spherical coordinates.

13. Representations of Lie Algebras: 4/28/14

Last time we discussed representations of SU_2 and SO_3: we saw that for every $n \geq 0$, there's a unique irreducible representation of SU_2 with dimension $n+1$ and character

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \longmapsto e^{in\theta} + e^{i(n-2)\theta} + \cdots + e^{-in\theta}.$$

This is realized on homogeneous polynomials on $\mathbb{C}^2$ of degree n.

Thus, for every *even* $n \geq 0$, there's a unique irreducible representation of SO_3 with character

$$\mathrm{rot}_\varphi \longmapsto e^{in\varphi/2} + e^{i(n-1)\varphi/2} + \cdots + e^{-in\varphi/2}.$$

This is because there's a map $\mathrm{SU}_2 \to \mathrm{SO}_3$ sending $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right) \mapsto \mathrm{rot}(2\theta)$, and the representations thus induced are realized on the space of harmonic polynomials on $\mathbb{R}^3$ of degree $n/2$ (i.e. those killed by the Euclidean Laplacian $\Delta = \partial_{xx} + \partial_{yy} + \partial_{zz}$).

In this lecture, we turn to a Lie-algebraic perspective, which is useful because many proofs in the general case involve embedding SU_2 into some other compact Lie group.

Recall that all representations are taken to be continuous (and therefore smooth) into some $\mathrm{GL}_N(\mathbb{C})$. If $V = \mathbb{C}^N$ is a representation of SU_2, then there's a map $\rho : \mathrm{SU}_2 \to \mathrm{GL}_N(\mathbb{C})$, which can be differentiated into a map $\mathrm{d}\rho : \mathrm{Lie}(\mathrm{SU}_2) \to M_N(\mathbb{C})$ of Lie algebras, sending $X \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\rho(e^{tX})$. $\mathrm{d}\rho$ is real linear, and therefore can be extended into a complex linear

$$\mathrm{d}\rho_{\mathbb{C}} : \underbrace{\mathrm{Lie}(\mathrm{SU}_2) \otimes \mathbb{C}}_{\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))} \to M_N(\mathbb{C}).$$

This is because after differentiating the conditions for SU_2, one has that

$$\begin{aligned} \mathrm{Lie}(\mathrm{SU}_2) &= \left\{ X \in M_2(\mathbb{C}) \mid X + \overline{X}^{\mathrm{T}} = 0, \mathrm{tr}(X) = 0 \right\} \\ &= \left\{ \begin{pmatrix} it & a+ib \\ -a+ib & -it \end{pmatrix} \mid a, b, t \in \mathbb{R} \right\}, \end{aligned}$$

which is a three-dimensional vector space, but

$$\begin{aligned} \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) &= \{ X \in M_2(\mathbb{C}) \mid \mathrm{tr}(X) = 0 \} \\ &= \left\{ \begin{pmatrix} t & \alpha \\ \beta & -t \end{pmatrix} \mid t, \alpha, \beta \in \mathbb{C} \right\}. \end{aligned}$$

This is three-dimensional over $\mathbb{C}$, so any real basis for $\mathrm{Lie}(\mathrm{SU}_2)$ is a complex basis for $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$, so $\mathrm{Lie}(\mathrm{SU}_2) \otimes \mathbb{C} = \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$.

Both $\mathrm{d}\rho$ and $\mathrm{d}\rho_{\mathbb{C}}$ respect the Lie algebra structure, i.e. they commute with the Lie bracket. In fact, ρ also extends to a representation $\rho_{\mathbb{C}} : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$, and we can regard $\mathrm{d}\rho_{\mathbb{C}}$ as its derivative. The argument can be by checking that differentiating and complexifying in either order produce the correct result, or noticing that the irreducible representations can be acted on in this way, and thus all of them can.

[26]To use this, it's necessary to check that the Laplacian is in fact surjective.

[27]The hat on top of the direct sum indicates the completed direct sum, i.e. taking the direct sum and then completing the corresponding space.

Now, one can compute $d\rho_{\mathbb{C}}$ for irreducible representations. A convenient basis for $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$ is

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{13.1}$$

These names are standard.

We'll compute how $d\rho_{\mathbb{C}}$ acts on h, e, and f. Then, an $\mathbb{R}$-basis for $\mathrm{Lie}(\mathrm{SU}_2)$ is given by $\langle ih, i(e+f), i(e-f)\rangle$, so one can get information about the representations.

Let V_n be the representation on homogeneous polynomials of degree n: if $P \in V_n$ and $g \in \mathrm{SU}_2$, then $gP(g\mathbf{x}) = P(\mathbf{x})$ for $\mathbf{x} \in \mathbb{C}^2$. A basis for V_n is given by $x^n, x^{n-1}y, x^{n-2}y^2, \ldots, y^n$, so we'll compute how each of these is acted upon. The action in question is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x^i y^j = (dx - by)^i (-cx + ay)^j;$$

This is clearer as a right action, but in general it's more confusing to mix left and right actions in the same place. Thus,

$$\begin{aligned}
(d\rho_{\mathbb{C}}(h)) \cdot (x^i y^j) &= \frac{d}{dt} \rho_C(\exp(th)) \cdot (x^i y^j) \bigg|_{t=0} \\
&= \frac{d}{dt} \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} x^i y^j \bigg|_{t=0} \\
&= \frac{d}{dt} e^{t(j-i)} (x^i y^j)|_{t=0} \bigg|_{t=0} \\
&= (j-i) x^i y^j,
\end{aligned}$$

which is where $(n-2i)$ comes from: in this basis, the $x^i y^j$ are eigenvectors, with eigenvalues $j-i$, corresponding to $n, n-2, \ldots, -n+2, -n$. Remember, though, $h \notin \mathrm{Lie}(\mathrm{SU}_2)$, just in $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$.

Since f is nilpotent, its exponential is really simple:

$$\begin{aligned}
d\rho_{\mathbb{C}}(f)(x^i y^j) &= \frac{d}{dt} \rho_{\mathbb{C}}(\exp(tf))(x^i y^j) \bigg|_{t=0} \\
&= \frac{d}{dt} \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} x^i y^j \bigg|_{t=0} \\
&= \frac{d}{dt} x^i (-tx + y)^j \bigg|_{t=0} \\
&= -j x^{i+1} y^{j-1}.
\end{aligned}$$

$d\rho_{\mathbb{C}}(f)$ maps each basis element to a multiple of the one before it, and e is essentially the same, but in the opposite direction. There's some scalar factor which doesn't end up mattering.[28]

$$\begin{tikzcd}
0 & x^n \arrow[l, bend right, "f"'] \arrow[r, bend right, "e"'] \arrow[loop below, "h", "-n"'] & x^{n-1}y \arrow[l, bend right, "f"'] \arrow[r, bend right, "e"'] \arrow[loop below, "h", "2-n"'] & \cdots \arrow[l, bend right, "f"'] \arrow[r, bend right, "e"'] & xy^{n-1} \arrow[l, bend right, "f"'] \arrow[r, bend right, "e"'] \arrow[loop below, "h", "n-2"'] & y^n \arrow[l, bend right, "f"'] \arrow[r, bend right, "e"'] \arrow[loop below, "h", "n"'] & 0
\end{tikzcd} \tag{13.2}$$

This is the Lie algebra structure. The basis elements satisfy the relations $[h,e] = 2e$, $[h,f] = 2f$, and $[e,f] = h$ (which you can check at home), and this is also true of their derivatives, so if one goes there and back again, it's possible to calculate the difference (e.g. $e \circ f$ versus $f \circ e$, the difference is by h).

We could have also used this to check irreducibility: if $U \subseteq V$ were invariant under SU_2, then it's also invariant under $\mathrm{Lie}(\mathrm{SU}_2)$ and thus under $\mathrm{Lie}(\mathrm{SU}_2) \otimes \mathbb{C}$ (since that just throws in complex scalars). Thus, U is invariant under $d\rho_{\mathbb{C}} e$, $d\rho_{\mathbb{C}} f$, and $d\rho_{\mathbb{C}} h$. Looking at the eigenspaces of h in particular, $x^i y^j \in U$ for some i and j (as long as U is nonzero), so by using e and f, this means all of them are in U, and thus $U = V_n$.

Since SU_2 is simply connected, then we know that every representation of Lie algebras (i.e. a Lie algebra homomorphism $\mathrm{Lie}(\mathrm{SU}_2) \to M_n(\mathbb{C})$, sort of a differentiated analogue of a group representation) must arise as the derivative of a representation of SU_2, and (nontrivial to prove) must be a sum of the irreducible representations shown.

[28]Except, of course, in positive characteristic…

Note that the commutation relations for h, e, and f force the following: if $hv = av$, then $v' = ev$ has the property that $hv' = (a+2)v'$, and $v'' = fv$ has $hv'' = (a-2)v''$. Thus, e is sometimes called the raising operator, as it raises the eigenvalue for h, and f is correspondingly called the lowering operator.

Looking Forward. We saw that in SU_2, everything is conjugate to a matrix of the form $r_\theta = \left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right)$, which is also conjugate to $r_{-\theta}$, and the irreducible representations are parameterized by $n \geq 0$ in $\mathbb{Z}$ with an explicit character formula. There was also an integration formula for class functions,

$$\int_{\mathrm{SU}_2} f = \frac{2}{\pi} \int_\theta f(r_\theta) \sin\theta \, \mathrm{d}\theta,$$

where the left side of the equation is integrated with respect to the Haar measure.

We also looked at U_n, in which everything is conjugate to a matrix of the form

$$r(\theta_1, \dots, \theta_n) = \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix},$$

so the eigenvalues are determined up to permutation; for any $\sigma \in S_n$, $r(\theta_1, \dots, \theta_n) \sim r(\theta_{\sigma(1)}, \dots, \theta_{\sigma(n)})$. Then, the irreducible representations are parameterized by n-tuples of integers $k_1 \geq k_2 \geq \cdots \geq k_n$, and the integration formula is

$$\int_{\mathrm{U}_n} f = \int_{(\theta_1, \dots, \theta_n)} f(r(\theta_1, \dots, \theta_n)) \Big| \prod_{j<k} \left(e^{i\theta_j} - e^{i\theta_k}\right) \Big| \, \mathrm{d}\theta.$$

These both can be generalized: suppose G is a compact, connected Lie group, such a U_n, SO_n, or $(S^1)^n$. There's a notion of a torus within G, a compact, connected, abelian Lie subgroup, which thus must be isomorphic to $(S^1)^n$, buy a proof found in the textbook.[29] Then, a maximal torus is a closed subgroup of G that is a torus and maximal with respect to these properties. It turns out that all maximal tori are conjugate, and that everything in G is conjugate to something in the maximal torus. Specifically, if $W = \text{Normalizer}(T)/T$, then $t \in T \sim w(t)$ for $w \in W$.

In U_n, $W = S_n$; in SU_2, $W = \mathbb{Z}/2$. W is a finite group acting on T, so the irreducible representations of G correspond to those of T, quotiented by W (i.e. taking W-orbits), and the irreducible representations of T are easy to discover, since it is abelian.

Another important fact is that G is almost determined by (T, W), in a sense that will be clarified. G is in fact fully determined by the normalizer of T, but this isn't useful.

Thus, the main players in the rest of the game are the torus and the finite group that acts on it.

14. Maximal Tori and Weyl Groups: 4/30/14

Definition 14.1. A torus is a connected abelian compact Lie group (and therefore isomorphic to $(S^1)^n$). In any Lie group G, a closed subgroup $T \subset G$ is called a maximal torus if it isn't contained in any other torus.

For the rest of the course, G will denote a compact, connected Lie group.

For an example, if $G = \mathrm{U}_n$, then the maximal torus is the subgroup of diagonal matrices.

Claim. Every element of G is be conjugate to an element of the maximal torus T.

This fact will imply that all maximal tori of a given Lie group are conjugate; in U_n, this says that all matrices are diagonal in their own eigenbases.

The Weyl group W is defined as $W = (\text{Normalizer}(T))/T$. In the case of U_n, $W = S_n$, since the normalizer of T consists of all permutations of elements of T.

For orthogonal groups, however, the parity of the dimension matters.

- In SO_{2n}, the maximal torus is given by elements with the block form

$$\begin{pmatrix} \operatorname{rot}\theta_1 & & \\ & \ddots & \\ & & \operatorname{rot}\theta_n \end{pmatrix},$$

[29] But we can already see that, since exp is a global homomorphism in this case, any torus must be a quotient of $\mathbb{R}^n$.

where $\operatorname{rot}\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$. Thus, T consists of n copies of S^1; it's clear that this is a torus, but not yet that it's maximal; soon, this will be much easier to check.

The Weyl group can permute these blocks, and since $\operatorname{rot}\theta \in \mathrm{O}(2)$, then $\mathrm{O}(2)^n$ normalizes T, and is in fact $\mathrm{O}(2) \cdot S^n$. Thus, the Weyl group is $(\mathrm{O}(2)^n \cdot S_n)^{\det=1}$, which is finite with size $2^{n-1}n!$

- In SO_{2n+1}, the setup is very similar, but not the same: T consists of elements of the form
$$\begin{pmatrix} \operatorname{rot}\theta_1 & & & \\ & \ddots & & \\ & & \operatorname{rot}\theta_n & \\ & & & 1 \end{pmatrix},$$
so that T is n copies of S^1 again, and the Weyl group is $W = \mathrm{O}(2)^n \cdot S_n/\mathrm{SO}_{2n+1}$, which has size $2^n n!$

If T is a torus, let $X^*(T)$ denote the set of its irreducible representations.[30] If $T = (S^1)^n$, then it's abelian, so all of the irreducible representations are one-dimensional, and thus $X^*(T) \xrightarrow{\sim} \mathbb{Z}^n$, given by
$$(m_1, \dots, m_n) \longmapsto \left((\theta_1, \dots, \theta_n) \longmapsto e^{i\sum_j \theta_j m_j}\right).$$
Thus, $X^*(T) \cong \mathbb{Z}^{\dim(T)}$.

Claim. Irreducible representations of G are parameterized by W-orbits on $X^*(T)$ (sometimes written $X^*(T)/W$ or X^*/W), and there is an explicit character formula.

For example, in the case of U_n, $\mathbb{Z}^n/S_n \leftrightarrow \{k_1 \geq k_2 \geq \cdots \geq k_n\}$, which is how the irreducible representations have been parameterized in class thus far. The general proof will be very similar to the case for U_n, once we prove the integration formula.

Proposition 14.2. *The pair (W, T) almost determine G.*

It's completely equivalent to think of this as $(W, X^*(T))$, a finite group acting on a Lie group. In some cases, the tori are abstractly isomorphic, but the action of W is different.

Definition 14.3. A group A acting on a vector space V is called a reflection group if it is generated by reflections, i.e. $r \in A$ such that $r^2 = 1$ and r fixes a subspace of codimension 1 (so that there's a basis for V in which r is diagonal, with diagonal entries only 1 and exactly one -1).

Claim. W acting on $X^* \otimes \mathbb{R}$ is a reflection group.

We will eventually prove this by embedding $\mathrm{SU}(2) \to G$, putting reflections inside G.

This is a very restrictive notion, because it's very hard to have a finite reflection group,[31] and turns the classification of compact Lie groups into a question of combinatorial geometry. Thus, one can classify all of the finite reflection groups arising from a compact Lie group G.

In the end, this will lead to a classification of compact Lie groups themselves, into S^1, U_n and SU_n, SO_{2n} and SO_{2n+1}, symplectic groups, and five exceptional groups, as well as products of these groups.

In U_n, W acting on X^* looks like $S_n \circlearrowright \mathbb{Z}^n$ by coordinate permutations, and is generated by transpositions (which can be thought of as the reflection through the hyperplane "between" the i- and j-axes), so it's a reflection group. In SO_{2n+1}, the action of $S_n \cdot \{\pm 1\}^n \circlearrowright \mathbb{Z}^n$ can act by coordinate permutation (the S_n bit) and changing signs (the other part), so there are more reflections, e.g. $x_i \leftrightarrow x_j$, $x_i \leftrightarrow -x_i$, and $x_i \leftrightarrow -x_j$. One can check that these generate the group.

What does "almost" mean in Proposition 14.2? Here's a simple example of its failure: $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$, given by maximal tori of the form $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \operatorname{rot}\theta & 0 \\ 0 & 1 \end{smallmatrix}\right)$ respectively, but in both cases have Weyl group $\mathbb{Z}/2$. Thus, for both, $X^*(T) \cong \mathbb{Z}$, and W acts by $m \mapsto -m$.

The map $\mathrm{SU}_2 \to \mathrm{SO}_3$ sending $r_\theta \mapsto \operatorname{rot} 2\theta$ wraps the torus upstairs twice around the one downstairs, and we would like some way to distinguish these.

Definition 14.4. Let T act on $\mathrm{Lie}(G)$ by the adjoint action; then, a root of T is a nonzero character $\chi \in X^*(T)$ that occurs in $\mathrm{Lie}(G) \otimes \mathbb{C}$. Then, $\Phi \subseteq X^*(T)$ denotes the set of roots.

---

[30] This notation comes from algebraic geometry, since this set is contravariant in T.

[31] What tends to happen is that two reflections generate an infinite number of elements, unless their angle is a rational multiple of π.

This will enable us to state the following much better result.

Claim. (W, X^*, Φ) determine G up to isomorphism. Moreover, W acting on X^* is a finite reflection group generated by reflections through the roots (i.e. there's a one-to-one correspondence): for each reflection r there's a unique (up to sign) root α in the -1-eigenspace.

In some sense, roots parameterize reflections, so it's possible to classify systems of the form (W, X^*, Φ) as a means to understanding Lie groups.

In U_n, the roots are the characters

$$\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} \longmapsto e^{i(\theta_j - \theta_k)},$$

corresponding to $(0, 0, \dots, 0, 1, 0, \dots, 0, -1, 0, \dots, 0)$, with a 1 in the j^{th} place and a -1 in the k^{th} place; geometrically, these correspond to reflections.[32] We can actually prove this right here: the adjoint of the maximal torus T of diagonal matrices with entries of the form $e^{i\theta_j}$ on $\mathrm{Lie}(\mathrm{U}_n) \otimes \mathbb{C} = M_n(\mathbb{C})$ is $\mathrm{Ad}(t) : X \mapsto tXt^{-1}$, for $X \in M_n(\mathbb{C})$, so $x_{jk} \mapsto e^{i\theta_j} e^{-i\theta_k} x_{jk}$. Thus, the $(j,k)^{\text{th}}$ entry is multiplied by $e^{i(\theta_j - \theta_k)}$, and these are exactly the roots.

In SO_{2n}, the roots can be written in $\mathbb{Z}^n$ (instead of $X^*(T)$) as $(0, \dots, 0, \pm 1, 0, \dots, 0, \pm 1, 0, \dots, 0)$, with the ± 1 in the j^{th} and k^{th} slots, or just $(0, \dots, 0, \pm 1, 0, \dots, 0)$, in just a single entry. These embody three kinds of reflections, though exactly how depends on sign. The roots refine the reflection.

The notion of a root feels pretty strange, e.g. there's no multiplicities. But they index reflections geometrically, so they're actually pretty natural.

Proposition 14.5. *All maximal tori of a given compact Lie group G are conjugate; thus, the maximal torus T of G is unique up to conjugacy.*

Proof. Fix $T \subseteq G$ a maximal torus, so that any $g \in G$ is conjugate to an element of T, so thus any two tori are conjugate. But why does this follow? In the classical, explicit case, one can pick eigenvectors, but this isn't true in general. The following trick is useful beyond the scope of this proof.

Let T' be another maximal torus, and find a $t' \in T'$ such that $\overline{\langle t' \rangle} = T'$ (i.e. $Z(t') = T'$). This can be done because $T = (S^1)^n$, so one can choose entries in these components that are linearly independent over $\mathbb{Q}$ and show that they have this property. This is known as a "generic element" of T'.

By assumption (everything in G is conjugate to something in T), there's a $g \in G$ such that $gt'g^{-1} \in T$, and thus $gT'g^{-1} \subseteq T$. But my maximality, this forces $gT'g^{-1} = T$. ⊠

15. Conjugacy of Maximal Tori: 5/2/14

Today, G is a compact connected Lie group and $T \subset G$ its maximal torus. Let $W = N(T)/T$ be its Weyl group and $X^*(T)$ be the set of irreducible representations of T, which is a free abelian group. Since all of the irreducible representations of T are one-dimensional, they can be identified with their characters, and thus the roots $\Phi \subset X^*(T)$. These are the characters occurring in $\mathrm{Lie}(G) \otimes \mathbb{C}$. Then, we aim to prove the following:

- the irreducible representations of G correspond to X^*/W,
- (W, X^*) is a reflection group, and
- (W, X^*, Φ) completely determines G.

For example, for U_n, T is the subgroup of diagonal matrices with diagonal entries $e^{i\theta_j}$ as usual, and $W = S_n$, acting on $X^*(T) \cong \mathbb{Z}^n$; the roots have the form $e_i \pm e_j$ when $i \neq j$.

There's an exceptional Lie group called G_2, which is 14-dimensional, but has a two-dimensional maximal torus T. Then, $W = D_{12}$ (the symmetries of the hexagon) and $X^* \cong \mathbb{Z}^2$. The roots, of which there are twelve, are plotted in Figure 2.

Similarly to how SO_n is the stabilizer of a symmetric bilinear form, Sp_n (the symplectic group) is the stabilizer for an alternating bilinear form.[33]

[32] The reflection through a vector is understood as the reflection through its normal hyperplane.

[33] Note that we don't look at stabilizers for higher-order forms because they're often trivial. For example, the space of alternating trilinear forms on $\mathbb{R}^n$ has dimension $\binom{n}{3}$, but $\mathrm{GL}_n(\mathbb{R})$ acts on this, and has dimension n^2. Speaking not rigorously, one would expect there to be a stabilizer if $\binom{n}{3} \leq n^2$, but this is only true when $n \leq 8$, and there is a nontrivial stabilizer in these cases. When $n = 8$, one gets $\mathrm{SU}_3 \cong \mathrm{SL}_3(\mathbb{R})$, and when $n = 6$,

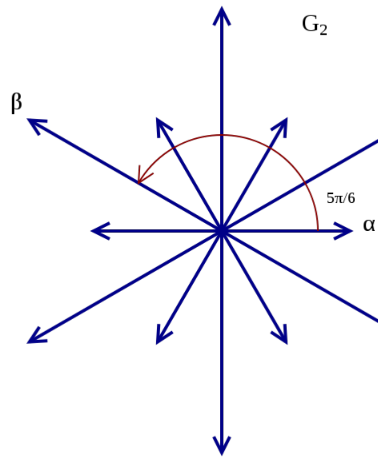

FIGURE 2. A depiction of the root system of the exceptional Lie group G_2. Source: http://en.wikipedia.org/wiki/G2_(mathematics).

Last time, we proved that any two maximal tori of G are conjugate, and it's also true (albeit a stronger fact) that any $g \in G$ is conjugate to an element of a given maximal torus, but that will be explained later. However, to finish the proof, we need to understand why, if t is a generic element of a maximal torus T' and T is another maximal torus, then there's a $g \in G$ such that $gt'g^{-1} \in T$. This will be a Lie algebra argument, setting the stage for what comes later.

Choose a generic $t' \in T'$, i.e. $\overline{\langle t' \rangle} = T'$ (in some sense, it needs to be irrational enough). Thus, $t' = \exp(X')$ for some $X' \in \mathrm{Lie}(T')$/ We'll show that there exists a $g \in G$ such that $\mathrm{Ad}(g)X' \in \mathrm{Lie}(T)$; then, by exponentiating, $gt'g^{-1} \in T$, so $gT'g^{-1} \subseteq T$, and thus they are equal, by the maximality of T.

Fix an inner product $\langle , \rangle$ on $\mathrm{Lie}(G)$ invariant by Ad; for example, when $G = \mathrm{U}_n$, $\mathrm{Lie}(G) \subseteq M_n(\mathbb{C})$, and a viable inner product is

$$\langle A, B \rangle = \sum_{i,j=1}^{n} A_{ij} \overline{B_{ij}}.$$

In the general case, such an inner product exists because G is compact, so it's possible to average. Then, choose a $t \in T$ such that $\overline{\langle t \rangle} = T$, and let $X \in \mathrm{Lie}(T)$ such that $\exp(X) = t$. Then, choose a $g \in G$ that minimizes the distance from $\mathrm{Ad}(g)X'$ to X under this inner product. That is, we're trying to minimize

$$\begin{aligned} \left\|\mathrm{Ad}(g)X' - X\right\|^2 &= \langle \mathrm{Ad}(g)X' - X, \mathrm{Ad}(g)X' - X \rangle \\ &= \left\|X'\right\|^2 + \|X\|^2 - 2\langle \mathrm{Ad}(g)X', X \rangle, \end{aligned}$$

since the inner product is Ad-invariant. Thus, the goal is to maximize $\langle \mathrm{Ad}(g)X', X \rangle$. A maximum exists by compactness, so assume it's attained at $g = e$, by replacing X' with $\mathrm{Ad}(g)X'$ (so that its derivative at e is zero). Eventually, this will imply that $X' \in \mathrm{Lie}(T)$, but this requires some explanation: basically, if this is conjugate, so was the original X'.

What is the derivative of this? We saw that $\frac{\mathrm{d}}{\mathrm{d}Y} \mathrm{ad}(g)X = [X, Y]$, so for all $Y \in \mathrm{Lie}(G)$,

$$\langle [Y, X], X' \rangle = 0, \tag{15.1}$$

The invariance of the inner product means that $\langle \mathrm{Ad}(g)X_1, \mathrm{Ad}(g)X_2 \rangle = \langle X_1, X_2 \rangle$, so after differentiating with respect to g, one obtains

$$\langle [Y, X_1], X_2 \rangle + \langle X_1, [Y, X_2] \rangle = 0, \tag{15.2}$$

i.e. $\mathrm{ad}(Y)$ is skew-symmetric. Thus, one can apply it to (15.1) to show that $\langle X', [Y, X] \rangle = 0$ for all $Y \in \mathrm{Lie}(G)$. (15.2) also implies that $\mathrm{ad}(Y)$ has imaginary eigenvalues, which we'll use later.

it's another classical group. But when $n = 7$, it's not a classical group, and is actually G_2. If one does this instead for symmetric "definite" forms, one gets groups called $\mathrm{SO}_{p,q}$, which are noncompact and therefore out of the scope of this class. $\mathrm{Lie}(\mathrm{SO}_{p,q}) \neq \mathrm{Lie}(\mathrm{SO}_n)$, but their complexifications are the same.

Geometrically, $X' \perp [X,Y]$ for all $Y \in \mathrm{Lie}(G)$, which forces it to also be perpendicular to the space L spanned by $[Y,X]$ for all Y. But $\overline{\langle \exp(X) \rangle} = T$, and we want to show that $X^{(\in)} \mathrm{Lie}(T)$ as well. We'll break this down into two parts.

Proposition 15.3.

(1) $\dim(L) = \dim(G) - \dim(T)$.
(2) *Then,* $L \perp \mathrm{Lie}(T)$.

Once these two pieces fall into place, then $L = (\mathrm{Lie}(T))^{\perp}$, so $X' \perp L$ and thus $X' \in \mathrm{Lie}(T)$.

Proof of Prop 15.3. For (1), it's enough to show that $\dim\{Y \mid [Y,X] = 0\} = \dim T$, since this is the null space of $[-,X]$. This, however, is the space of things that commute with X, which is $\mathrm{Lie}(T)$, because if $[Y,X] = 0$, then $\exp(TY)$ commutes with $\exp(tX)$ and thus also with T; then, the statement follows by the maximality of T.

For (2), we wish to show that $\langle [Y,X], \mathrm{Lie}(T) \rangle = 0$; since the Lie bracket is skew-symmetric, then this becomes $\langle -[X,Y], \mathrm{Lie}(T) \rangle = \langle Y, [X, \mathrm{Lie}(T)] \rangle$, which is 0 because $\mathrm{Lie}(T)$ is abelian, ⊠

Thus, the proof of the conjugacy of maximal tori follows. It's a bit involved, but (15.2) comes up in other proofs.

The next thing we'll address is this business about the Weyl group W being a reflection group, generated by reflections (that is, in its action on X^*, or on the corresponding vector space $X^* \otimes \mathbb{R}$) by reflections s_α corresponding to roots $\alpha \in \Phi$.

Given an inner product space $(H, \langle , \rangle)$ and a vector $x \in H$, we can define the reflection through x to be the reflection through the hyperplane $x^{\perp}$ that negates x, i.e.

$$r_x(w) = w - \frac{2\langle w, x \rangle}{\langle x, x \rangle} x.$$

Thus, this acts as the identity on $x^{\perp}$, but sends $x \mapsto -x$.

The inner product on $\mathrm{Lie}(G)$ gives an inner product on everything else. In particular, it induces one on $X^* \otimes \mathbb{R}$, so defines the reflections that negate the roots α. Any G-invariant inner product will do, though.

Since $\mathrm{Lie}(T) \subseteq \mathrm{Lie}(G)$, then the inner product on $\mathrm{Lie}(G)$ induces one on $\mathrm{Lie}(T)$, and since $X^*(T)$ is the set of characters $T \to \mathbb{C}^*$, then the map $\chi \mapsto (1/i)\,\mathrm{d}\chi$ identifies $X^* \otimes \mathbb{R} \xrightarrow{\sim} \mathrm{Lie}(T)^*$ (the dual space $\mathrm{Hom}(\mathrm{Lie}(T), \mathbb{R})$), because $\mathrm{d}\chi$ is purely imaginary. Thus, the inner product on $\mathrm{Lie}(T)$ induces one on $\mathrm{Lie}(T)^*$ as normally, and thus on $X^* \otimes \mathbb{R}$.

Now, we want to produce a reflection $s_\alpha \in W$ for each root $\alpha \in \Phi$. We'll actually do much more, making a homomorphism $\iota_\alpha : \mathrm{SU}(2) \to G$, and then build the reflections as

$$\iota_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \longmapsto s_\alpha.$$

The abstract proof will be given next time, but until then, here is an example in the case $G = \mathrm{U}_3$. The roots are $\alpha = (1,-1,0)$, $(-1,1,0)$, $\beta = (1,0,-1)$, $(-1,0,1)$, $\gamma = (0,1,-1)$, and $(0,-1,1)$. They sit inside $\mathbb{Z}^3$ naturally identified with the character group of the torus, $X^*(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3})$, so

$$\iota_\alpha \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ -\overline{b} & \overline{a} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\iota_\beta \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -\overline{b} & 0 & \overline{a} \end{pmatrix}$$

$$\iota_\gamma \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & -\overline{b} & \overline{a} \end{pmatrix}.$$

Thus, we can use what we already know about $\mathrm{SU}(2)$ to learn things about the representations of other Lie groups.

Once again, let G be a connected, compact Lie group with maximal torus T. Let $X^*(T)$ be its character group, acted on by $W = N(T)/T$, the Weyl group. The roots $\Phi \subseteq X^*(T)$ are the chracters that occur in the adjoint action of $\mathrm{Lie}(G)\otimes\mathbb{C}$. Then, (X^*, W, Φ) will determine G, and so forth.

Today, though, we'll continue showing that (X^*, W) is a reflection group by producing enough reflections to classify it.

Take, as last time,a G-invariant inner product on $\mathrm{Lie}(G)$, which induces a G-invariant inner product on $\mathrm{Lie}(T)$, and thus also $X^*(T)\otimes\mathbb{R} \xrightarrow{\sim} \mathrm{Lie}(T)^*$ sending $\chi \mapsto (1/i)\,\mathrm{d}\chi$. This space of real functionals on T also gets an inner product.

Today, we'll show that for every root α, the reflection through the hyperplane orthogonal to α, i.e.

$$s_\alpha(v) = v - \frac{2\langle v, \alpha\rangle}{\langle \alpha, \alpha\rangle}\alpha \in W.$$

Formally, s_α is a reflection of the real vector space $X^*(T)\otimes\mathbb{R}$ (i.e. an automorphism), so we really mean that there's an element of W inducing it; then, uniqueness of that element follows straightforwardly. In fact, the representation $W \to \mathrm{Aut}(X^*\otimes\mathbb{R})$ is faithful, so each $w \in W$ is determine by its action on X^*. This representation is given as follows: if $\chi \in X^*$, $t \in T$, and $w \in W$, then $w \cdot t = wtw^{-1}$, and $(w\cdot\chi)(w\cdot t) = \chi(t)$, so that $w\cdot\chi(t) = \chi(w^{-1}t)$.

For example, if $G = \mathrm{U}_n$, then $T = \left\{\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n}\end{pmatrix}\right\}$ as usual, so the character group is $X^* \cong \mathbb{Z}^n$ via

$$(m_1, \dots, m_n) \longleftrightarrow e^{i\sum m_j\theta_j} = \chi.$$

The Lie algebra of T is $\mathrm{Lie}(T) = \left\{\begin{pmatrix} ia_1 & & \\ & \ddots & \\ & & ia_n\end{pmatrix}\right\}$, so

$$\mathrm{d}\chi : \begin{pmatrix} ia_1 & & \\ & \ddots & \\ & & ia_n\end{pmatrix} \longmapsto i\sum_{j=1}^n m_j a_j$$

is an *imaginary*-valued functional, so when one divides it by i, it becomes real-valued, and $\mathbb{Z}^n\otimes\mathbb{R}$ gives the full $\mathbb{R}^n$ of functionals.

Constructing the U_n-invariant inner product shows that on $X^*\otimes\mathbb{R}$,

$$\|(m_1, \dots, m_n)\| \propto \sum m_j^2,$$

where we don't really care about the constant of proportionality, since this is only used to define orthogonality for reflections.

The roots of U_n are $e_i - e_j \in \mathbb{Z}^n$, where $e_1, \dots, e_n$ is the standard basis, and thus $s_{e_i - e_j}$ swaps e_i and e_j and leaves everything else fixed. They Weyl group is isomorphic to S_n, and with this context, $d_{e_i - e_j}$ is just the transposition $(i\ j)$.

There are very few reflection groups in the world, so this is the crucial step in the classification of compact Lie groups.

Lemma 16.1. *Write $\mathfrak{g} = \mathrm{Lie}(G)\otimes\mathbb{C}$, and for any character $\chi \in X^*(T)$, write $\mathfrak{g}_\chi$ for the X^*-eigenspace, i.e.*

$$\mathfrak{g}_\chi = \{X \in \mathfrak{g} \mid \mathrm{Ad}(t)X = \chi(t)X\},$$

so that $\mathfrak{g}_\chi \neq 0$ iff $\chi \in \Phi\cup\{0\}$. Then,

(1) $\mathfrak{g}_0 = \mathrm{Lie}(T)\otimes\mathbb{C}$,
(2) $\overline{\mathfrak{g}_\alpha} = \mathfrak{g}_{-\alpha}$, *and*
(3) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

The proof of this lemma will mostly be a matter of definition checking.

In the case of U_n and $\alpha = e_i - e_j$, let $A_{i,j}$ be the matrix with a 1 in the $(i,j)^{\text{th}}$ position and 0 elsewhere. Then, $\mathfrak{g}_\alpha = \mathbb{C}\cdot A_{i,j}$.[34]

[34]In general, these spaces will be one-dimensional, but we haven't proven that yet.

Proof of Lemma 16.1.

- For (1), clearly $\mathrm{Lie}(T) \subseteq \mathfrak{g}_0$, since $\mathfrak{g}_0$ is the set of things in $\mathfrak{g}$ that commute with T, and equality is by maximality of T (if not, one can use exp to pass back and forth and find a contradiction).
- For (2), an element of $\mathfrak{g}$ is of the form $X + iY$ for $X, Y \in \mathrm{Lie}(G)$, so the conjugate is $\overline{X + iY} = X - iY$.[35] If $a \in \mathrm{Lie}(G)$, the real Lie algebra, then it's conjugate with itself.

  For example, if $G = \mathrm{U}_n$,
  $$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \frac{i}{2}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$
  and both of the matrices on the right are in $\mathrm{Lie}(\mathrm{U}_n)$. Thus, the conjugate switches them, so one obtains $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$.

  Returning to the proof, suppose that $X \in \mathfrak{g}_\alpha$, so that $\mathrm{Ad}(T)X = \alpha(t)X$ for any $t \in T$. Let's conjugate that:
  $$\begin{aligned} \overline{\mathrm{Ad}(T)X} &= \overline{\alpha(t)}\overline{X} \\ &= \alpha(t)^{-1}\overline{X}, \end{aligned}$$
  because α is a character of a compact group, so $\alpha^{-1} = \overline{\alpha}$. This is akin to the unit circle. Then,
  $$\Longrightarrow \mathrm{Ad}(t)\overline{X} = \alpha(t)^{-1}\overline{X},$$
  because Ad commutes with conjugation. Thus, $\overline{X} \in \mathfrak{g}_{\alpha^{-1}}$.
- For (3), take $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$. Conjugation respects the Lie bracket, so
  $$\begin{aligned} \mathrm{Ad}(t)[X, Y] &= [\mathrm{Ad}(t)X, \mathrm{Ad}(t)Y] \\ &= [\alpha(t)X, \alpha(t)Y] \\ &= \alpha\beta(t)[X, Y]. \end{aligned}$$
  Note that $\alpha\beta$ isn't always a root, but $\mathfrak{g}_{\alpha\beta}$ still makes sense. ⊠

Recall that by differentiating the condition on the invariant inner product of G, one obtains the skew-symmetric property (15.2), and thus that $\mathrm{ad}(X)$ acts as a skew-symmetric transform. Also, recall that $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$ is the complex span $\langle h, e, f \rangle$, where h, e, and f are given in (13.1); specifically, that $[h, e] = -2e$, $[h, f] = -2f$, and $[e, f] = h$.

Proposition 16.2. *Let α be a root and $X \in \mathfrak{g}_\alpha$ be nonzero. Then, there exists a map $\iota_\alpha : \mathrm{SU}(2) \to G$, and therefore also its derivative $\mathrm{d}\iota_\alpha : \mathrm{Lie}(\mathrm{SU}_2) \to \mathrm{Lie}(G)$ and its complexification $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) \to \mathfrak{g} = \mathrm{Lie}(G) \otimes \mathbb{C}$ such that:*
$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} &\longmapsto \lambda X \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &\longmapsto -\lambda\overline{X} \end{aligned}$$
for some $\lambda \neq 0$, and such that $\iota_\alpha(\mathrm{SU}_2)$ commutes with $\ker(\alpha) \subset T$.

This $\mathfrak{g}_\alpha$ is called the root space, and once this proposition is proven,
$$s_\alpha = \iota_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
ends up being the desired reflection.

For example, when $G = \mathrm{U}_3$ and $\alpha = e_1 - e_3$, the map will send $\mathrm{SU}(2)$ into U_3 via the first and third entries:
$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \longmapsto \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -\overline{b} & 0 & \overline{a} \end{pmatrix}.$$

[35] Be warned: this is not the standard conjugation operation when G is a complex matrix group.

Proof of Proposition 16.2. We'll construct the maps in reverse order; since SU_2 is simply connected, we can pull $\mathrm{d}\iota_\alpha$ up to ι_α.

To give $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) \to \mathfrak{g}$, we need to product $h', e', f' \in \mathfrak{g}$ such that $[h', e'] = 2e'$, $[h', f'] = 2f'$, and $[e', f'] = h'$. Then, we can send $e \mapsto e'$, $f \mapsto f'$, and $h \mapsto h'$. We'll set $e' = X$ and $f' = -\overline{X}$, so that $h' = -[X, \overline{X}]$ is forced. Now, we have to check the commutator relations.

We know $\mathrm{Ad}(t)e' = \alpha(t)e'$, so for $Y \in \mathrm{Lie}(T)$, we ca differentiate in the Y-direction. Thus, for $\alpha \in (\mathrm{Lie}(T))^*$, $\mathrm{ad}(Y)e' = i\alpha(Y)e'$, and therefore $[Y, e'] = i\alpha(Y)e'$.

Since $X \in \mathfrak{g}_\alpha$ and $\overline{X} \in \mathfrak{g}_\alpha$, then $h' \in \mathfrak{g}_0 = \mathrm{Lie}(T)_\mathbb{C}$. Thus, $[h', e'] = i\alpha(h')e'$, and similarly, $[h', f'] = -i\alpha(h')f'$, so it's enough to show that $\alpha(h')$ is a positive real divided by i. In particular, it is possible to scale X and $\overline{X}$ accordingly so that these brackets go to 2 (this is where λ comes from). This is the only thing we need to check, so in some sense, there are no choices!

Note that $\overline{h} = -[\overline{X}, X] = -h'$, so $h'/i \in \mathrm{Lie}(G)$, and thus $\alpha(h'/i) \in \mathbb{R}$, so we just need to check that it's negative. We'll show that it's nonzero, but a more careful analysis of the signs in the same line of reasoning will prove that it is in fact negative.

If $h' = 0$, then X and $\overline{X}$ commute with each other, and also with $\ker(\alpha)$ (since they act by $\pm\alpha$), so one obtain a torus of dimension at least $\dim(\ker(\alpha)) + 2 = \dim(T) + 1$. This argument is a little tricky because of complexification, but one can pick elements that span the same real vector space.

So we've shown that $[X, \overline{X}] \neq 0$, but we want $\alpha([X, \overline{X}])$ to be nonzero. Divide by i, so that we're in the real Lie algebra. In $\mathrm{Lie}(T)$, $(1/i)[X, \overline{X}]$ is in the α-direction, i.e., if we use the inner product to identify $\mathrm{Lie}(T)$ with $(\mathrm{Lie}(T))^*$, then this lies parallel to α. We'll verify this next time, though; stay tuned.

17. Mapping SU_2 into Compact Lie Groups: 5/7/14

As usual, let G denote a compact, connected Lie group, $T \subset G$ be its maximal torus, $X^*(T)$ be its group of characters, and $W = N(T)/T$.

Last time, we were showing that if $\mathfrak{g} = \mathrm{Lie}(G)_\mathbb{C}$ and $\mathfrak{g}_\alpha = \{v \in G \mid \mathrm{Ad}(t)v = \alpha(t)v\}$ for roots $\alpha \in \Phi$, then if $X \in \mathfrak{g}_\alpha \setminus 0$, then there's a homomorphism $\iota_\alpha : \mathrm{SU}(2) \to G$ such that $\mathrm{d}\iota_\alpha$ sends

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \longmapsto \lambda_1 X \qquad \text{and} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \longmapsto \lambda_2 \overline{X},$$

for $\lambda_1, \lambda_2 \neq 0$. These matrices lie in $\mathrm{Lie}(\mathrm{SU}_2))_\mathbb{C} = \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$.

For example, if $G = \mathrm{U}(3)$ and $\alpha = (-1, 0, 1)$, then $\mathfrak{g}_\alpha = \mathbb{C}\left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Then, $\ker(\alpha)$ is the things killed by its action:

$$\ker(\alpha) = \begin{pmatrix} z_1 & 0 & 0 \\ 0 & z_2 & 0 \\ 0 & 0 & z_1 \end{pmatrix},$$

so

$$\iota_\alpha : \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \longmapsto \begin{pmatrix} a & 0 & b \\ 0 & a & 0 \\ -\overline{b} & 0 & \overline{a} \end{pmatrix}.$$

Last time, we constructed a map $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) \to \mathfrak{g}$ such that $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \mapsto e'$, which is a multiple of X; $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \mapsto f'$, a multiple of $\overline{X}$, and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \mapsto h' = [e', f'] \propto [X, \overline{X}]$. But it still remains to see why $\alpha(h') \neq 0$, so that these multiples are nonzero. We showed that $h' \neq 0$, but since $h' \propto [X, \overline{X}]$ and $\overline{h'} = -h'$, then h' must be imaginary, and can be written as $h' = iH$ for some $H \in \mathrm{Lie}(G)$. Since $h' \neq 0$, then $H \neq 0$ too.

What needs to still be shown is that H lies in the direction of α, i.e. that $H \perp \ker(\alpha) \subseteq \mathrm{Lie}(T)$, or that $H \in (\mathrm{Lie}(T))^*$ is proportional to α (the identification between $\mathrm{Lie}(T)$ and its dual made via the invariant inner product).

Take $Y \in \ker(\alpha) \subseteq \mathrm{Lie}(T)$; this is somewhat sloppy notation, as we've used α both as the character $T \to \mathbb{C}^*$, and also its derivative $\mathrm{Lie}(T) \to \mathbb{R}$. Then, the goal is to show that $Y \perp H$. Well, $H \propto [X + \overline{X}, (X - \overline{X})/i]$, then since $Y \in \ker(\alpha)$, then $Y(X + \overline{X}) = Y((X - \overline{X})/i) = 0$. Thus, what we want follows from the invariance of the inner product:

$$\left\langle Y, \left[X + \overline{X}, \frac{X - \overline{X}}{i}\right]\right\rangle = -\left\langle [X + \overline{X}, Y], \cdots \right\rangle = 0.$$

For example, in the case U_3 as before, $H = \begin{pmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{pmatrix}$, so it's in the same direction as α.

Thus, there's a Lie algebra homomorphism $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) \to \mathrm{Lie}(G) \otimes \mathbb{C}$. But $\mathrm{Lie}(\mathrm{SU}_2) \subseteq \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$ is spanned by

$$\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \right\},$$

which are respectively sent to some scalar multiples of $\{X + \overline{X}, (X - \overline{X})/i, ih'\}$. Thus, under this homomorphism,

$$\begin{tikzcd} \mathrm{Lie}(\mathrm{SU}_2) \arrow[r, "\varphi"] \arrow[d, hook] & \mathrm{Lie}(G) \arrow[d, hook] \\ \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C})) \arrow[r] & \mathrm{Lie}(G) \otimes \mathbb{C}. \end{tikzcd}$$

The map constructed above is thus the $\mathbb{C}$-linear extension of φ, but since SU_2 is simply connected, then φ comes from some $\iota_\alpha : \mathrm{SU}_2 \to G$.

The textbook uses a different strategy that can be generalized (e.g. to noncompact Lie groups): the goal is for SU_2 to commute with $\ker(\alpha) \subset T$, so look at $Z(\ker(\alpha))$: this is a compact Lie group containing $\ker(\alpha)$ centrally.

Definition 17.1. The rank of a compact Lie group G is the dimension of its maximal torus.

This notion (which we could have provided several lectures ago) is well-defined because the maximal tori are conjugate.

Thus, $Z(\ker(\alpha)/\ker(\alpha)$ is a compact group of rank 1, though it might not be connected. However, we know that the only compact, connected groups of rank 1 are S^1, SU(2), and $\mathrm{SO}(3) = \mathrm{SU}(2)/Z(\mathrm{SU}(2))$. Thus, excluding S^1, there's a map $\mathrm{SU}(2) \to Z(\ker(\alpha))/\ker(\alpha)$, which lifts (via the Lie algebra) to $\iota_\alpha : \mathrm{SU}(2) \to Z(\ker(\alpha)) \subset G$.

This makes it clearer where SU(2) comes from: it's the only interesting group of rank 1, and it's easy to take rank-1 slices of root spaces.

Reflections. With all of these copies of SU_2 floating around, it's possible to say interesting things about reflections. Define

$$s_v : x \longmapsto x - \frac{2\langle x, v\rangle}{\langle v, v\rangle} v$$

to be the reflection through $v^\perp$, and also define $v^\vee$ ("v-check") as $v^\vee = 2v/\langle v, v\rangle$.

Theorem 17.2. *For every $\alpha \in \Phi$, the element $\iota_\alpha\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ normalizes T, and its class $s_\alpha \in W$ acts on* $\mathrm{Lie}(T)^*$ *as a reflection through $\alpha^\perp$, i.e. $\iota_\alpha\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = s_\alpha$, so every $\beta \in \Phi$ can be reflected: $s_\alpha(\beta) = \beta - \langle \beta, \alpha^\vee\rangle \alpha \in \Phi$, and moreover $\langle \beta, \alpha^\vee\rangle \in \mathbb{Z}$.*

It's also true that these s_α generate W, but that's not immediately important, and we'll prove it when we get there.

The condition that $\langle \beta, \alpha^\vee\rangle \in \mathbb{Z}$ is actually even more restrictive: it forces $\cos^2 \theta_{\alpha\beta} \in \frac{1}{4}\mathbb{Z}$, severely limiting the number of possible reflection groups that arise from Lie groups.

Proof of Theorem 17.2. $\iota_\alpha\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ centralizes $\ker(\alpha) \subset T$ (or $\ker(\alpha) \subseteq \mathrm{Lie}(T)$), and $\iota_\alpha\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ negates $H \in \mathrm{Lie}(T)$, as

$$\mathrm{Ad}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] = -\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is in the α-direction. Thus, since these generate T, then $\iota_\alpha\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ normalizes T, so it must act as a reflection through $\alpha^\perp$ on $\mathrm{Lie}(T)^*$.

That $s_\alpha(\beta)$ is a root if β is comes from the fact that W preserves roots.

Lastly, to show the integrality of $\langle \beta, \alpha^\vee\rangle$, it happens to be true that $\iota_\alpha\left(\begin{smallmatrix} z & 0 \\ 0 & z^{-1} \end{smallmatrix}\right)$ acts on $\mathfrak{g}_\beta$ by the character $z \mapsto z^{\langle \beta, \alpha^\vee\rangle}$, but this is a character of S^1, and thus must be an integer. ⊠

Returning to $G = \mathrm{U}(3)$, $\alpha = (-1, 0, 1)$, and so forth, let $\beta = (1, -1, 0)$, so that $\alpha^\vee = 2\alpha/\langle \alpha, \alpha\rangle = \alpha$, and thus $\langle \alpha^\vee, \beta\rangle = 1$, which says that

$$\iota_\alpha\begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} = \begin{pmatrix} z & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^{-1} \end{pmatrix},$$

and β applied to that is just dividing z by 1, yielding z, which we could have calculated explicitly.

At the level of the Lie algebra, the exponent must be proportional to $\langle\alpha,\beta\rangle$, because ι_α centralizes $\ker(\alpha)$. To check which multiple it actually is, compute with $\beta=\alpha$ $\iota_\alpha\left(\begin{smallmatrix} z & 0 \\ 0 & z^{-1}\end{smallmatrix}\right)$ acts on $\mathfrak{g}_\alpha$ be z^2, so

$$\mathrm{Ad}\begin{pmatrix} z & 0 \\ 0 & z^{-1}\end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}=\begin{pmatrix} 0 & z^2 \\ 0 & 0\end{pmatrix},$$

so the exponent is $2\langle\alpha,\beta\rangle/\langle\alpha,\alpha\rangle=\langle\alpha^\vee,\beta\rangle$ (so that it's 2 when $\beta=\alpha$, giving the action of z^2).

Thus, not only is $\langle\beta,\alpha^\vee\rangle\in\mathbb{Z}$, but its specific value indicates how α and β interact. Again, it's extremely restrictive: it means that the projection of β onto $\mathbb{R}\alpha$ is a multiple of $\alpha/2$. This also holds if one switches α and β, because

$$\langle\alpha,\beta^\vee\rangle\langle\beta,\alpha^\vee\rangle=4\frac{\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\leq 4 \tag{17.3}$$

by the Cauchy-Schwarz inequality, Thus, $\langle\alpha,\beta^\vee\rangle\in\{-4,\dots,4\}\subset\mathbb{Z}$, and is only ± 4 if equality holds in (17.3), i.e. when α and β are proportional. Also, if $\theta_{\alpha\beta}$ is the angle between them, then

$$\cos^2\theta_{\alpha\beta}=\frac{\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle},$$

so it must lie within $\frac{1}{4}\mathbb{Z}$, drastically limiting the number of possible choices:

$$\theta_{\alpha\beta}\in\{0^\circ,30^\circ,45^\circ,60^\circ,90^\circ,120^\circ,135^\circ,150^\circ,180^\circ\}.$$

This restriction is akin to forcing α to preserve a lattice.

18. Classical Compact Groups: 5/9/14

As usual, the *dramatis personæ* are a compact connected Lie group G, its maximal torus T, the Weyl group W, the irreducible representations X^*, and the roots Φ.

Last time, we showed that for every $\alpha\in\Phi$, there's a reflection $s_\alpha\in W$ negating α, and $\langle\beta,\alpha^\vee\rangle\in\mathbb{Z}$ for all $\alpha,\beta\in\Phi$.

In order to provide basic examples of what we've been talking about, consider the classical compact Lie groups:

- $\mathrm{O}(n)$ is the group of $\mathbb{R}$-linear automorphisms of $\mathbb{R}^n$ preserving the form $\sum x_i^2$.
- $\mathrm{U}(n)$ is the group of $\mathbb{C}$-linear automorphisms of $\mathbb{C}^n$ preserving the Hermitian form $\sum x_i\overline{x_i}$.

There's one more thing that can be added here, given by the quaternions

$$\mathbb{H}=\{a+bi+cj+dk\mid a,b,c,d\in\mathbb{R}, ij=k, jk=i, ki=j, i^2=j^2k^2=-1\}.$$

There are many ways to put $\mathbb{C}\hookrightarrow\mathbb{H}$ as rings, but it's conventional to send $a+bi\mapsto a+bi$.

The quaternions have a conjugation, which behaves a lot like that of $\mathbb{C}$:

$$\overline{a+bi+cj+dk}=a-bi-cj-dk.$$

Thus, there's a Lie group $\mathrm{Sp}(2n)$,[36] which is the group of $\mathbb{H}$-linear automorphisms preserving the form $\sum x_i\overline{x_i}$, where the conjugation is the one just introduced. Since $\mathbb{H}$ isn't commutative, then it's necessary to specify that these automorphisms be left $\mathbb{H}$-linear.

These three examples are the basic examples: everything else, even the sporadic groups, will come in some sense from them.

As another example of the symplectic group, $\mathrm{Sp}(2)$ is the set of $f:\mathbb{H}\to\mathbb{H}$ such that f commutes with left multiplication and $f(x)f(\overline{x})=x\overline{x}$. But if f commutes with left multiplication, then it's given by right multiplication: $f(x)=xq$, forcing $q\overline{q}=1$. Thus, $\mathrm{Sp}(2)\cong\{q\in\mathbb{H}\mid q\overline{q}=1\}\cong\mathrm{SU}(2)$. These can be thought of as quaternionic matrices acting on the right.

These three families of groups are all compact because the forms are positive definite on real vector spaces, so they're contained within some larger orthgonal group $\mathrm{O}(2n)$ or $\mathrm{O}(4n)$, which is also compact.

Each of these comes with a standard complex representation: $\mathrm{O}(n)$ acts on $\mathbb{R}^n$ and thus $\mathbb{C}^n$, $\mathrm{U}(n)$ acts on $\mathbb{C}^n$, and $\mathrm{Sp}(2n)$ acts on $\mathbb{H}^n\cong\mathbb{C}^{2n}$.[37] These will be called the standard representations. It's also possible to form these with $\mathrm{SO}(n)$ and $\mathrm{SU}(n)$, but they're not as important in this realization.

$\mathrm{O}(m)$ and, if m is even, $\mathrm{Sp}(m)$ play an important role in real and quaternionic representations.

[36]Some sources call this group $\mathrm{Sp}(n)$, so be careful about the dimension of $\mathbb{H}^n$ being used.

[37]The isomorphism as as left $\mathbb{C}$-vector spaces.

Proposition 18.1. *If G is a compact group and V is an irreducible m-dimensional representation of G, then the following are equivalent:*

(1) $\chi_V(g) \in \mathbb{R}$.
(2) $V \cong V^*$.
(3) $V \cong \overline{V}$.[38]
(4) *G preserves either a symplectic or skew-symmetric bilinear form $V \times V \to \mathbb{C}$.*
(5) *G acting on V is conjugate to either a subgroup of $\mathrm{O}(m)$ (if it preserves the symmetric form) or $\mathrm{Sp}(m)$ (if it preserves the skew-symmetric form).*
(6) *G acting on V comes from either a representation over $\mathbb{R}$ (in the symmetric case), or over $\mathbb{H}$ (in the skew-symmetric case).*

This implies that if m is odd, then G must preserve a symmetric form, and thus is conjugate to a subgroup of $\mathrm{O}(m)$ and comes from a real representation.

This large amount of statements can be distilled into saying that if a representation is dual to itself, in the sense of (2) or (3), then it preserves a symmetric or skew-symmetric form, and then so on. Thus, it is possible to think of quaternionic representations as complex representations that preserve a skew-symmetric form.

For example, D_8 and $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\} \subseteq \mathbb{H}^\times$ have the same character table and both have a unique two-dimensional representation (the reflections and rotations of the square), and thus are self-dual (since the dual representation must have the same dimension). For D_8 it has a symmetric bilinear form, and can be conjugated into O(2), and comes from a real representation. But for Q_8, it comes from an $\mathbb{H}$-rep, and can be conjugated into Sp(2), as it preserve a skew-symmetric form.

This is a standard counterexample to the naïve idea that one could reconstruct a group from its characters, but it turns out that the category of representations of a group G (i.e., along with the data of the morphisms) is enough to reconstruct G.

Proof sketch of Proposition 18.1. If V^* is the dual representation of G, then $\chi_{V^*}(g) = \chi_V(g^{-1}) = \overline{\chi_V(g)}$, so if $\chi_V \in \mathbb{R}$, then $\chi_{V^*} = \chi_V = \chi_{\overline{V}}$, and $\overline{V}$ is the same vector space, but with scalar multiplication conjugated, which implies the equivalence of (1), (2), and (3).

For (3) $\iff$ (6), $V \cong \overline{V}$ iff there's a conjugate linear $J : V \to V$ (i.e. $J(\lambda v) = \overline{\lambda} J(v)$) commuting with G. Then, $J^2 : V \to V$ is linear, so by Schur's lemma, it's a scalar, and by looking at its trace, it's a real scalar. Now, there are two possibilities:

- If $J^2 > 0$, then λJ has the same property for any $\lambda \in \mathbb{C}$ (except $\lambda = 0$), so without loss of generality, let $J^2 = 1$. Now, it's possible to extract a real representation: $\mathbb{R}^n = \mathrm{fix}(J)$ (there is a lot to show here, but it ends up working out).
- If $J^2 < 0$, then λJ has the same property again when $\lambda \neq 0$, so one can assume that $\lambda^2 = -1$. Then, J corresponds to the action of $j \in \mathbb{H}$, so the actions of i and j generate an $\mathbb{H}$-action.

There are plenty of details to fill in here, but not too many headaches.

To get up to (5), take a quadratic (or Hermitian) form and average it; then, the result will end up in $\mathrm{O}(m)$ (or $\mathrm{Sp}(m)$). This is the same argument that shows that every representation over $\mathbb{C}$ preserves some Hermitian form; it's clear why $\mathrm{O}(n)$ preserves the standard dot product, but less clear why $\mathrm{Sp}(m)$ preserves a symplectic (i.e. $\mathbb{C}$-linear, nondegenerate, and skew-symmetric) form on $\mathbb{C}^n$, where $m = 2n$.

To see this, observe that $\mathrm{Sp}(n) \circlearrowright \mathbb{H}^n$ by definition preserves $\sum x_i \overline{x_i}$; thus, it also preserves the form $h(x,y) = \sum x_i \overline{y_i}$, in the sense that $h(gx, gy) = h(x,y)$. This is still $\mathbb{H}$-valued, though, so take the imaginary part $\mathrm{Im}\, h : \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{C}$, which is skew-symmetric. The notation is unfortunate, but makes sense: $\mathrm{Im}(a + bi + cj + dk) = (cj + dk)j \in \mathbb{C}$. Then, $\mathrm{Im}(h)$ is a skew-symmetric, complex linear form. ⊠

Returning to the maximal tori, all of our examples can be collected into a table (not all of which we've formally proved, but none of which is particularly hard to show). In Table 1, some of the notation needs to be explained. The rotation matrix is simply $\mathrm{rot}(\theta_i) = \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix}$, and $S_n \cdot (\pm 1)^{n-1} \circlearrowright \mathbb{Z}^n$ is the action given by the semidirect product $S_n \rtimes (\mathbb{Z}/2)^{n-1}$, and so on.

Notice that (X^*, W) is isomorphic for $\mathrm{SO}(2n+1)$ and $\mathrm{Sp}(2n)$; this implies that they have a lot in common (e.g. similar representations), despite having different dimensions and parities.

[38] If one thinks of these as matrix representations, then V^* sends $A \mapsto A^{\mathrm{T}}$ and $\overline{V}$ sends $A \mapsto \overline{A}$.

G	T	W	Φ
$\mathrm{U}(n)$	$\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix}$	S_n	$e_i - e_j$
$\mathrm{SO}(2n)$	$\begin{pmatrix} \mathrm{rot}\,\theta_1 & & \\ & \ddots & \\ & & \mathrm{rot}\,\theta_n \end{pmatrix}$	$S_n \cdot (\pm 1)^{n-1} \circlearrowright \mathbb{Z}^n$	$\pm e_i \pm e_j$
$\mathrm{SO}(2n+1)$	$\begin{pmatrix} \mathrm{rot}\,\theta_1 & & & \\ & \ddots & & \\ & & \mathrm{rot}\,\theta_n & \\ & & & 1 \end{pmatrix}$	$S_n \cdot (\pm 1)^{n} \circlearrowright \mathbb{Z}^n$	$\pm e_i \pm e_j, \pm e_j$
$\mathrm{Sp}(2n)$	$\begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix}$	$S_n \cdot (\pm 1)^{n}$	$\pm e_i \pm e_j, \pm 2e_j$
$\mathrm{SU}(n)$	$\left\{ \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} \mid \sum \theta_i = 0 \right\}$	S_n	$e_i - e_j$

TABLE 1. Summary of information for some classical compact Lie groups.

Next time, the properties of these roots will be abstracted away into the notion of a root system.

19. ROOT SYSTEMS: 5/12/14

Recall the setup from last time, specifically Table 1. In all cases, an inner product on $X^* \cong \mathbb{Z}^n$ can be taken to be proportional to $\sum x_i^2$. Also notice that the dimension of the maximal torus of $\mathrm{SU}(n)$ is one less than that of $\mathrm{U}(n)$, and its irreducible characters are given by $X^* \cong \mathbb{Z}^n / \langle (1, \dots, 1) \rangle$, via

$$(m_1, \dots, m_n) \longmapsto e^{i(m_1\theta_1 + \cdots + m_n\theta_n)}.$$

Now, we can draw some pictures of the root systems, as in Figure 3. In each case, $\dim T = 2$, and we're viewing the roots as within $X^* \otimes \mathbb{R}$.

Notice that the root systems have some similarities: after rotation and scaling, those of $\mathrm{SO}(5)$ and $\mathrm{Sp}(4)$ look very similar, as do those of $\mathrm{SO}(4)$ and $\mathrm{SU}(2) \times \mathrm{SU}(2)$. Since these root systems essentially determine the group, then there are homomorphisms $\mathrm{Sp}(4) \to \mathrm{SO}(5)$ and $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. These are nearly isomorphisms; in fact, they are isogenies, i.e. surjections with finite kernel, which induce isomorphisms on Lie algebras. They're also universal covers.

One can read lots of other stuff off too, e.g. that the copy of the root system of $\mathrm{SU}(3)$ inside G_2 also induces a homomorphism.

Looking a bit deeper into these exceptional isogenies, we have $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ and $\mathrm{Sp}(4) \to \mathrm{SO}(5)$ as above, and also saw a simpler example earlier on in the class: that of $\mathrm{SU}(2) \to \mathrm{SO}(3)$. Finally, there's one last example, $\mathrm{SU}(4) \to \mathrm{SO}(6)$. The way we wrote this stuff down makes it surprising that this data determines the group, but here are some illustrations as to why it happens.

- For $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$, we saw that $\mathrm{SU}(2) \cong \mathbb{H}^{(1)}$ (i.e. the unit quaternions), but $\mathbb{H}^{(1)} \times \mathbb{H}^{(1)}$ acts on $\mathbb{H}$ from the left and right, and preserves $x\overline{x} = a^2 + b^2 + c^2 + d^2$. Thus, there's a map $\mathrm{SU}(2) \times \mathrm{SU}(2)$ to the space of $\mathbb{H}$-automorphisms preserving this quadratic form, which is $\mathrm{SO}(4)$. Then, one can check that it's an isomorphism on Lie algebras.

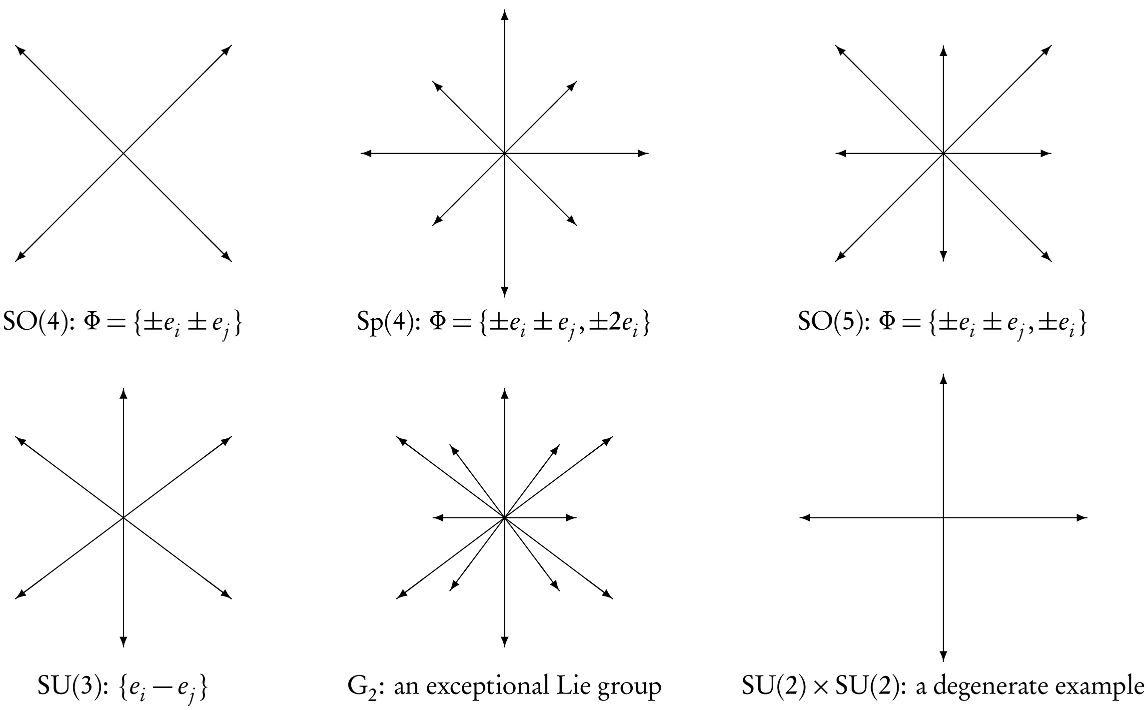

FIGURE 3. The root systems of several compact Lie groups with two-dimensional maximal tori. For G_2, notice the outer and inner hexagons, and for $\mathrm{SU}(2) \times \mathrm{SU}(2)$, each pair of vectors comes from one of the factors, since the roots of SU(2) are given by $\pm\alpha$ for a single α.

Another way to view this is that $\mathrm{SU}(2) \circlearrowright \mathbb{C}^2$, so $\mathrm{SU}(2) \times \mathrm{SU}(2) \circlearrowright \mathbb{C}^2 \otimes \mathbb{C}^2$. The representation of SU(2) on $\mathbb{C}^2$ is self-dual, because it's the only one of its dimension, and so the product representation is also self-dual. Thus, it preserves a symmetric bilinear form (since $\mathrm{SU}(2) \circlearrowright \mathbb{C}$ preserves a symplectic form), so it must be within SO(4), as proved in Proposition 18.1.

- SU(4) acts on $V = \mathbb{C}^4$, so we want to make a six-dimensional representation from V, and hope that it's self-dual and preserves a quadratic form. The trick is to always pick one side and then try to make a representation from it. Six dimensions suggests $\wedge^2 V$, which has dimension $\binom{4}{2} = 6$. But why is it self-dual? This is trickier: V isn't self-dual, since its character isn't real-valued (and thus changes under conjugation), but

$$\chi_{\wedge^2 V} = e^{i(\theta_1+\theta_2)} + e^{i(\theta_1+\theta_3)} + \cdots + e^{i(\theta_3+\theta_4)},$$

and since the θ_i must sum to zero, then after conjugating, $e^{-i(\theta_1+\theta_2)} = e^{i(\theta_3+\theta_4)}$, so this character is preserved under conjugation, and thus $\wedge^2 V$ is self-dual.

Next, we need a bilinear form $\wedge^2 V \times \wedge^2 V \to \mathbb{C}$. How about something like $(\omega, \omega') \mapsto \omega \wedge \omega' \in \lambda^4 V$, which is one-dimensional? Thus, SU(4) acts trivially, and the form itself is symmetric.

When we discuss spin representations, we'll be able to construct a map going in the other direction.
- $\mathrm{Sp}(4) \to \mathrm{SO}(5)$ is similar, but a bit trickier, and is worth working through.

The general idea here is that there are several useful examples to convince yourself that isomorphic root systems do imply isomorphic groups, as it's clearer in these smaller-rank cases.

We have now arrived at the point where we can state one of the major classification theorems of the course.

Theorem 19.1 (Classification theorem, first version). *If G is a compact, connected Lie group, then G is determined up to isomorphism by (X^*, W, Φ).*

Recall that here, X^* is a free abelian group, W is a finite group acting on X^*, and $\Phi \subseteq X^*$. In fact, though, G is already almost determined by (X^*, W).

Theorem 19.2 (Classification theorem, second version)**.** *G is determined up to isogeny by $(X^* \otimes \mathbb{R}, W, \Phi)$; that is, if G and G' have the same triple $(X^* \otimes \mathbb{R}, W, \Phi)$, then* $\mathrm{Lie}(G) \cong \mathrm{Lie}(G')$, *or (equivalently) they have the same universal covering.*

This is particularly handy for the Lie-algebraic perspective: on this level, for example, there is no difference between Lie(Sp(4)) and Lie(SO(5)).

Along with the theorems presented above, there's also an axiomatic way of presenting this classification. Unfortunately, there are several competing, but slightly different, definitions of root systems. Here's the one from the textbook, which was influenced by Serre, who was in turn influenced by Bourbaki.

Definition 19.3. A root system is a tuple (V, W, Φ), where V is a finite-dimensional real vector space, W is a finite group, and $\Phi \subset V$, such that:

(1) for every $\alpha \in \Phi$, there's a reflection $s_\alpha \in W$ such that $s_\alpha : \alpha \mapsto -\alpha$, and that W is generated by such reflections.[39]
(2) $\langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$ for any $\alpha, \beta \in \Phi$.

There are also two somewhat minor conditions:

(3) Φ spans V, and
(4) if $\alpha, \beta \in \Phi$ are proportional, then $\alpha = \pm\beta$.

If (3) is omitted, there is no great consequence: the theory is changed slightly, and becomes more inconvenient. Thus, it is often kept for convenience. Furthermore, axiom (4) is true for compact Lie groups, but we haven't shown that yet.

There's again a less precise version up to isogeny, which is sufficient for most applications, and a more precise one, as follows.

Definition 19.4. A root datum is a triple $(L, \Phi, \Phi^\vee)$, such that L is a finite abelian group, $\Phi \subset L$, and $\Phi^\vee \subset L^* = \mathrm{Hom}(L, \mathbb{Z})$, satisfying similar axioms.

This notion won't be as important.

Remark (Dual groups). One amazing consequence of the more precise formulation is that, along with Theorem 19.1, the map $(L, \Phi, \Phi^\vee) \mapsto (L^*, \Phi^\vee, \Phi)$ gives an involution on isomorphism classes of compact Lie groups. This is still a bit of a miracle and not very well-understood (there's no satisfactory construction; it's all done after the classification). Here are some examples:

$$\begin{aligned} \mathrm{SO}(2n+1) &\longleftrightarrow \mathrm{Sp}(2n) \\ \mathrm{SU}(n) &\longleftrightarrow \mathrm{U}(n)/Z(\mathrm{U}(n)) \\ \mathrm{SU}(2) &\longleftrightarrow \mathrm{SO}(3). \end{aligned}$$ ◁

Now, Theorem 19.2 says that G is determined up to isogeny by its root system. There are two more important facts that we'll also need to address.

- Every root system occurs thus (well, up to a minor nuance that we'll address).
- Every root system is a direct sum of root systems of the classical groups SU(n), SO(n), and Sp(n), and the five exceptional groups G_2, F_4, E_6, E_7, and E_8.

20. Exceptional Root Systems: 5/14/14

Recall that a root system (V, Φ, W) is a combination of a real vector space V, $\Phi \subset V$, and a finite group W acting on V and preserving Φ, subject to the following axioms.

(1) For each $\alpha \in \Phi$, there's a reflection $s_\alpha \in W$ negating α, and these generate W. s_α is given by $v \mapsto v - \langle v, \alpha^\vee \rangle \alpha$, with $\alpha^\vee \in V^*$.
(2) For all $\alpha, \beta \in \Phi$, $\langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$.
(3) Φ spans V. This axiom is less important.
(4) If $\alpha, \beta \in \Phi$ are proportional, then $\alpha = \pm\beta$, $\alpha = \pm 2\beta$, or $\alpha = \pm\frac{1}{2}\beta$. This axiom is less important.

[39] This s_α is of the form $s_\alpha(v) = v - \langle v, \alpha^\vee \rangle \alpha$, with $\alpha^\vee \in V^*$. In this sense, a reflection differs from the identity by a rank-1 term in the direction of α.

If G is a compact, connected Lie group, then $(X^*(T)\otimes\mathbb{R},\Phi,W)$, where W is its Weyl group, satisfies all of these axioms (though we have yet to show (4) and that the roots generate W for (1)), except for (3). Thus, $(\mathrm{span}_{\mathbb{R}}(\Phi),\Phi,W)$ is a root system. Thus triple $(X^*(T)\otimes\mathbb{R},\Phi,W)$ determines G up to isogeny (i.e. isomorphism on Lie algebras).

There is a direct sum on root systems, given by that which induces the direct product on groups:

$$(V_1,\Phi_1,W_1)\oplus(V_2,\Phi_2,W_2)=(V_1\oplus V_2,(\Phi_1,0)\cup(0,\Phi_2),W_1\times W_2).$$

With this, the following facts are true:

- Every root system is a direct sum of classical root systems and those given by G_2, F_4, E_6, E_7, and E_8.
- If one ignores axiom (3), every root system is a direct sum of the root systems mentioned above and the trivial root system $(\mathbb{R},\Phi=\emptyset,W=\{1\})$.
- Whether or not axiom (3) is used, every root system arises from a compact Lie group.

Of course, these are all theorems! There's a lot to show here.

Interestingly, Killing discovered some of the exceptional Lie groups by discovering extra root systems, which ended up corresponding to these groups.

		$X^*\otimes\mathbb{R}$	Φ	W
A_n	$\mathrm{SU}(n)$	$\mathbb{R}^n/\langle(1,\dots,1)\rangle$	e_i-e_j	S_n
B_n	$\mathrm{SO}(2n+1)$	$\mathbb{R}^n$	$\pm e_i\pm e_j,\pm e_i$	$S_n\cdot(\mathbb{Z}/2)^n$
C_n	$\mathrm{Sp}(2n)$	$\mathbb{R}^n$	$\pm e_i\pm e_j,\pm 2e_i$	$S_n\cdot(\mathbb{Z}/2)^n$
D_n	$\mathrm{SO}(2n)$	$\mathbb{R}^n$	$\pm e_i\pm e_j$	$S_n\cdot(\mathbb{Z}/2)^{n-1}$

TABLE 2. The classical root systems. Often, they're given the names $\mathrm{A}_n,\dots,\mathrm{D}_n$.

Another definition of root systems adds an inner product to V and doesn't need the Weyl group.

Definition 20.1 (Alternate)**.** A root system is a a tuple $(V,\Phi\subseteq V)$ such that for each $\alpha\in\Phi$, $s_\alpha(\Phi)=\Phi$, where s_α is the orthogonal reflection $\alpha\mapsto-\alpha$, i.e.

$$s_\alpha(v)=v-\frac{2\langle\alpha,v\rangle}{\langle\alpha,\alpha\rangle}\alpha,$$

and subject to the same axioms (2), (3), and (4) as in the previous definition.

This definition is less natural: a W-invariant inner product certainly exists, since W is a finite group, but it might not be unique. In the other direction, given this definition of a root system, the Weyl group can be recovered as $W=\langle s_\alpha\rangle$. The notion of isomorphism of root systems is the same in this framework, but one must be mindful that isomorphisms do not always preserve the inner product.

Exceptional Groups. In order to construct the exceptional groups, one somehow must use exceptional root systems. One way to do this involve Dynkin diagrams, but that's less important and is left for the text.

Suppose L is a lattice, i.e. a finitely generated abelian group together with an inner product $L\times L\to\mathbb{Z}$ (i.e. extending $\mathbb{Z}^n\to\mathbb{R}^n$ gives an inner product). It's not really important right now that it's $\mathbb{Z}$-valued.

From L, one can build a root system Φ_L (that lacks condition (3)), in which

$$\Phi_L=\{\alpha\in L\mid s_\alpha L=L\}\subseteq L.$$

Specifically, we require the α to be primitive, so that if $\alpha=n\alpha'$ for some $n\in\mathbb{Z}$, then we replace α with α'. The condition in building the roots is that the lattice is reflected into itself. Then, the root system is given by $(L\otimes\mathbb{R},\Phi_L)$.

Since α is primitive, then $v-\langle v^\vee,\alpha^\vee\rangle\alpha\in L$ for all $v\in L$ is equivalent to requiring $\langle v,\alpha^\vee\rangle\in\mathbb{Z}$. That is,

$$\Phi_L=\left\{\alpha\mid\alpha^\vee=\frac{2\alpha}{\langle\alpha,\alpha\rangle}\in L^*\right\},$$

where

$$L^*=\{\lambda\in L\otimes\mathbb{R}\mid\langle\lambda,L\rangle\subseteq\mathbb{Z}\}.$$

In some cases, this is easier to remember, or clearer, than the standard definition.

For example, if $L=\mathbb{Z}^n$ with the standard inner product, then $\langle\alpha^\vee,\alpha^\vee\rangle=4/\langle\alpha,\alpha\rangle$, so $\langle\alpha,\alpha\rangle\in\{1,2,4\}$, but can't actually be 4 (because then, $\alpha^\vee=\alpha/2\notin L$, since L has to be primitive). Thus, $\langle\alpha,\alpha\rangle=1$, so $\alpha=\pm e_i$, or $\langle\alpha,\alpha\rangle=2$, and $\alpha=\pm e_i\pm e_j$. Thus, we get the root system B_n!

For any compact Lie group G, we can apply this to $L = X^*(T)$ together with an invariant inner product on $X^* \otimes \mathbb{R}$ (which can always be arranged to be integral). In most cases, you just get back the root system of G, but in two cases, you get something bigger: an exceptional root system. In this way we can construct G_2 and F_4.

For SU(3), the maximal torus is

$$\left\{ \begin{pmatrix} e^{i\theta_1} & & \\ & e^{i\theta_2} & \\ & & e^{i\theta_3} \end{pmatrix} \mid \theta_1 + \theta_2 + \theta_3 = 0 \right\},$$

so any three integers give a character, but if they're the same, then it's the trivial character:

$$(m_1, m_2, m_3) \longmapsto e^{i\sum_{j=1}^3 m_j \theta_j}.$$

Thus, $X^*(T) = \mathbb{Z}^3/\langle(1,1,1)\rangle$, which can be given an inner product by projecting the standard inner product on $\mathbb{Z}^3$ to $(1,1,1)$. This yields a hexagonal lattice Φ_2 such that $|\Phi_2| = 12$, and thus one obtains G_2. Maybe this isn't so interesting, because one can get it by hand, but it also induces an embedding $\mathrm{SO}_3 \to \mathrm{G}_2$ (via the external hexagon in the roots; the internal hexagon only comes up over characteristic 3!).

To get F_4, start with SO(8). Let $L = \mathbb{Z}-\mathrm{span}\{\pm e_i \pm e_j \mid 1 \le i, j \le 4\} \subset \mathbb{Z}^4$.[40] Thus, this is a four-dimensional checkerboard lattice, in that every other point of the full lattice $\mathbb{Z}^4$ is missing: certainly, $\pm\varepsilon_i \pm e_j \in L$, since the reflections just permute them around, but there are others, and thus

$$L^* = \mathbb{Z}^4 \cup \left(\mathbb{Z}^4 + \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)\right).$$

Then, for any $\beta \in L^*$, $\langle \beta, \beta \rangle \in \mathbb{Z}$, so if $\alpha^\vee \in L^*$, $\langle \alpha, \alpha \rangle \in \{1, 2, 4\}$ by the same reasoning. Thus, $\alpha \in \{\pm e_i \pm e_j, \pm e_1 \pm e_2 \pm e_3 \pm e_4\}$. For $\alpha = \pm e_1 \pm e_2 \pm e_3 \pm e_4$, then $\alpha^\vee = \alpha/2 \in L^*$, since we specifically included those. This is a root system of size 40, called F_4, and there's a map $\mathrm{SO}(8) \to \mathrm{F}_4$.

Finally, there's an eight-dimensional lattice that induces E_8, and E_6 and E_7 can be created by cutting down E_8. Start with the checkerboard lattice L in $\mathbb{Z}^8$ (created by throwing away every other point). That is,

$$L^* = \mathbb{Z}^8 \cup \left(\mathbb{Z}^8 + \left(\frac{1}{2}, \dots, \frac{1}{2}\right)\right).$$

There are multiple ways of extending this lattice L into a self-dual lattice Λ (i.e. $\Lambda = \Lambda^*$); right now, it has index 4 in its dual. Instead, we can take the E_8-lattice $\lambda = L \cup (L + (1/2, \dots, 1/2))$, which is a checkerboard along with a shifted checkerboard. This is self-dual, and $\langle x, x \rangle$ is even for all $x \in L$. A lattice with these two properties is called even unimodular lattice (i.e. determinant 1) lattice, and these exist only in dimensions $8k$, with the E_8 lattice the unique eight-dimensional one. (This construction also works in dimension $8k$ for other k).

To understand Φ_Λ, $\langle \alpha^\vee, \alpha^\vee \rangle = 4/\langle \alpha, \alpha \rangle \in \Lambda^* = \Lambda$, so $\langle \alpha^\vee, \alpha^\vee \rangle$ is one of 2 or 4, and thus $\langle \alpha, \alpha \rangle$ is either 1 or 2. But all of the lengths are even, so the lengths must be 2. Thus, $\Phi_\Lambda = \{\pm e_i \pm e_j, \pm(1/2)e_1 \pm (1/2)e_2 \pm \cdots \pm (1/2)e_8$ with an even number of $+$ signs$\}$. There are 112 vectors of the first type and 120 of the second type, so the E_8 root system has 240 roots, and thus $\dim(\mathrm{E}_8) = 248$ (the extra eight dimensions come from $\mathrm{Lie}(T)$).

21. Chambers: 5/16/14

Once again, the actors are:

- a compact, connected Lie group G,
- its maximal torus T,
- the roots $\Phi \subset X^*(T)$,
- and the Weyl group W.

We have mostly proven that $(X^*(T) \otimes \mathbb{R}, \Phi, W)$ forms a root system, and stated without proof that this determines G. Then, we discussed the classification of root systems.

Next, we'll show that the irreducible representations of G correspond to W-orbits of X^*, but before that, we'll need to check a few odds and ends. Recall that for $\alpha \in \Phi$, $\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid \mathrm{Ad}(t)X = \alpha(t)X$ for all $t \in T\}$, so that $\mathfrak{g}_0 = \mathrm{Lie}(T) \otimes \mathbb{C}$.

Proposition 21.1.

(1) *If α is a root, then* $\dim \mathfrak{g}_\alpha = 1$.

[40] Alternatively, this is just $\{(x_1, x_2, x_3, x_4) \mid \sum x_i$ is even$\}$.

(2) *If $\alpha, \beta \in \Phi$ are proportional, then $\alpha = \pm\beta$.*
(3) *W is generated by the reflections s_α for $\alpha \in \Phi$.*

It was possible to check these a while ago, but they weren't really needed until now. (3) develops ideas that we will use later.

Proof of Proposition 21.1. For (1), look at $\iota_\alpha : \mathrm{SU}(2) \to G$. Let $\mathfrak{g} = \mathrm{Lie}(G) \otimes \mathbb{C}$ as usual (intuitively, we want to complexify this because we're thinking about complex representations). Since G acts on $\mathfrak{g}$, then SU(2) does as well, and its action is trivial on ker(α on Lie(T)). Moreover, by the properties of α stated last time it was discussed, $\left(\begin{smallmatrix} z & 0 \\ 0 & z^{-1} \end{smallmatrix}\right)$ acts on $\mathfrak{g}_\beta$ by $\langle \beta, \alpha^\vee \rangle$.

Let

$$V = \bigoplus_{j \in \mathbb{Z}} \mathfrak{g}_{j\alpha} \subseteq \mathfrak{g},$$

so that V is preserved by $\iota_\alpha(\mathrm{SU}_2)$: $\mathrm{Lie}(\iota_\alpha(\mathrm{SU}_2)) = \left\langle X, \overline{X}, H \right\rangle$, with $X \in \mathfrak{g}_\alpha$, $\overline{X} \in \mathfrak{g}_{-\alpha}$, and $H \in \mathrm{Lie}(T)$. Recall also that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

	z^{-2}	z^{-1}	1	z^1	z^2
trivial	0	0	1	0	0
2-dim.	0	1	0	1	0
3-dim.	1	0	1	0	1

TABLE 3. A depiction of some irreducible SU_2-representations. For example, on $\langle x^2, xy, y^2 \rangle$, the diagonal matrix with entries (z, z^{-1}) acts on x^2 by z^2, xy by 1, and y^2 by z^{-2}.

Recall the general picture of irreducible SU_2-representations, illustrated in Table 3. If W is any representation of SU_2, then let

$$W_m = \left\{ w \in W \mid \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} w = z^m w \right\},$$

then $\dim W_0 = (\dim W^{\mathrm{SU}_2}) + \dim W_2$. This means that if there's a 1 in the z^2-term, then there's a 1 in the middle.

For this representation V, the decomposition is $\mathfrak{g}_{-2\alpha} \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{2\alpha}$, upon which $\left(\begin{smallmatrix} z & 0 \\ 0 & z^{-1} \end{smallmatrix}\right)$ acts respectively as $(z^{-4}, z^{-2}, 1, z^2, z^4)$, since $\iota_\alpha\left(\begin{smallmatrix} z & 0 \\ 0 & z^{-1} \end{smallmatrix}\right)$ acts on $\mathfrak{g}_\beta$ by $z\langle \alpha^\vee, \beta \rangle$. Thus, $\dim T = \dim V_0 \geq \dim V^{\mathrm{SU}_2} + \dim \mathfrak{g}_\alpha$, so $\dim \mathfrak{g}_\alpha \leq 1$ and $\mathfrak{g}_{k\alpha} = 0$ when $|k| \geq 2$ (by similar reasoning).

This nearly implies (2), too! Suppose α and β are proportional roots and suppose β is longer in a G-invariant inner product, so that $\beta = c\alpha$; if $\beta \neq \pm\alpha$, then $|c| > 1$. Then, since $\langle \alpha, \beta^\vee \rangle = 2\langle \alpha, \beta \rangle / \langle \beta, \beta \rangle = 2/c \in \mathbb{Z}$, then $c = \pm 2$, but we just saw that $\mathfrak{g}_{\pm 2\alpha} = 0$. Oops. Thus, $c = \pm 1$. ⊠

This last axiom isn't so important: relaxing it adds only one root system to the classification, which doesn't even show up in the compact case.

For (3), we want to make W act on something and make arguments based on this action. Thus, this group-theoretic argument turns into a geometric one. In fact, geometry is how you prove everything about the Weyl group.

Definition 21.2. Given a root system $(X^* \otimes \mathbb{R}, \Phi, W)$, a chamber is a connected component of $X^* \otimes \mathbb{R}$ minus the reflection hyperplanes $\mathbb{R}_\alpha^\perp$ (for each s_α).

The proof will boil down to an action of W on the chambers.[41]

TODO: add a picture of a chamber of SU_3.

In general, the number of chambers will be shown to be $|W|$, and the action of W is transitive and doesn't fix anything.

Claim.

(4) $\langle s_\alpha \rangle$ acts transitively, so that $|\langle s_\alpha \rangle|$ is at least the number of chambers, and
(5) no element of W save for e fixes any chamber (so that $|W|$ is at most the number of chambers).

Putting these together shows that $\langle W \rangle$ is equal to the number of chambers, which is equal to $|\langle s_\alpha \rangle|$, and that W acts simply transitively on the chambers (i.e. the action is transitive, and doesn't fix things in the sense given above).

[41] I guess that makes W a chamber group?

Proof of the claim. For (4), here's a way to get from one chamber to another: pick a point P in the first chamber and Q in the second such that the line between them doesn't go through the origin. Then, draw the line between them, which intersects some reflection lines in order; then, these reflections in that order bring P into Q. TODO: add a picture of this.

For (5), let $w \in W$, $w \neq e$, so that if w fixes a chamber C, then there's a $v \in C$ fixed by w (take any $v' \in C$ and average it under $\langle w \rangle$). We need to use compactness somehow, so translate this into group-theoretic terms: $v \in X^*(T) \otimes \mathbb{R} = \mathrm{Lie}(T)^* \cong \mathrm{Lie}(T)$ (via a G-invariant inner product on $\mathrm{Lie}(G)$). $\langle v, \alpha \rangle \neq 0$, since the vectors for which this is true were removed to form the chambers, so let $X \in \mathrm{Lie}(T)$ be the element corresponding to v, so that W fixes X.

Let $\widetilde{w} \in N(T)$ (that is, the normalizer) be a representative of w, since $W = N(T)/T$, and let $= \overline{\exp(\mathbb{R}x)} \subseteq T$. Then, S is compact, because T is, and it's connected, since a copy of $\mathbb{R}$ is dense inside of it. It's also an abelian group, so S is a torus, though it may not be maximal. Thus, $\widetilde{w}$ commutes with S, so let Z be the centralizer of S in G, which we will prove later is connected.

Specifically, to prove the claim, we'll need to assume that the centralizer of a torus is connected. This is slightly stronger than the notion of conjugacy of maximal tori...

With this assumption, we now can write that $\mathrm{Lie}(Z) = \{Y \in \mathrm{Lie}(G) \mid [X, Y] = 0\}$ (i.e. Y commutes with X), and since $X \in \mathrm{Lie}(T)$, then this is just $\mathrm{Lie}(Z) = \mathrm{Lie}(T)$, because v isn't perpendicular to any α. Thus, $Z = T$, so $\widetilde{w} \in T$, and thus $w = e$. ⊠

We've now shown completely that $(X^* \otimes \mathbb{R}, \Phi, W)$ is a root system. But while we're on the subject of chambers, each chamber gives rise to a basis (sometimes called a base, or a "system of simple roots") in Φ, i.e. a set $\{\alpha_1, \dots, \alpha_n\} \subseteq \Phi$ such that

(1) they are an $\mathbb{R}$-basis for the root system $\mathbb{R}\Phi$, and
(2) every root can be written $\sum n_i \alpha_i$, where $n_i \in \mathbb{Z}$ and are all either all ≥ 0 or ≤ 0 (so there's nothing of the form $\alpha_1 - \alpha_2$).

In particular, this divides the roots into $\Phi = \Phi^+ \cup \Phi^-$, those where al of the n_i are at least 0 or at most 0, respectively. These can be constructed by taking some chamber C and a $\lambda \in C$. Then, λ^* divides the space into two: $\{v \mid \langle \lambda, v \rangle \geq 0\}$ and $\{v \mid \langle \lambda, v \rangle \leq 0\}$. But no root lies along the line $\langle \lambda, v \rangle = 0\}$, since that's a reflection plane, so it lies in no chamber, and thus $\Phi = \Phi^+ \sqcup \Phi^-$, where $\Phi^+ = \{\langle \alpha, \lambda \rangle \geq 0\}$ and $\Phi^- = \{\langle \alpha, \lambda \rangle \leq 0\}$.

Then, it's possible to obtain a basis by choosing $\alpha_1, \dots, \alpha_n$ to be a set of minimal elements in Φ^+, i.e. those that cannot be written as a sum of two other roots in Φ^+. The proof, however, will have to be deferred to the next lecture.

22. Representations as Orbits of the Weyl Group: 5/19/14

The next step in the classification of Lie groups will be to prove that irreducible representations of a compact connected Lie group G correspond to W-orbits of $X^*(T)$.

Recall that a chamber in $V = X^*(T) \otimes \mathbb{R}$ is a connected component of

$$V - \bigcup_{\alpha \in \Phi} \alpha^\perp,$$

where $\alpha^\perp$ is the reflection hyperplane for a given root α, and that W acts simply transitively on the chambers.

For example, on U_n, $X^*(T) = \mathbb{Z}_n$ and $\Phi \supset \{e_i - e_j\}$; then, the reflection hyerplanes are $\alpha^\perp = \{x_i - x_j\}$. Thus, one chamber is the region $\{x_1 > x_2 > \cdots > x_n\} \subseteq \mathbb{R}^n$, and the others are $\{x_{\sigma(1)} > x_{\sigma(2)} > \cdots > x_{\sigma(n)}\} \subseteq \mathbb{R}^n$, for some $\sigma \in S_n$. It's clear in this case that the Weyl group S_n acts simply transitively on the chambers.

Given a chamber C, there's a division $\Phi = \Phi^+ \cup \Phi^-$, where $\Phi^+ = \{\alpha \mid \langle \alpha, \lambda \rangle \geq 0\}$ for any $\lambda \in C$ (they all lead to the same division), and $\Phi^- = \{\alpha \mid \langle \alpha, \lambda \rangle \leq 0\}$. Then, $\alpha \in \Phi^+$ iff $-\alpha \in \Phi^-$, and if $\alpha, \beta \in \Phi^+$ and $\alpha + \beta$ is still a root, then $\alpha + \beta \in \Phi^+$.

Here's a picture in the case $G = \mathrm{SU}_3$: TODO.

Definition 22.1. Given this division $\Phi = \Phi^+ \cup \Phi^-$, a system of simple roots for Φ is a set of minimal elements of Φ^+: they cannot be written as a sum of two other roots.

Example 22.2.

(1) $B = \{\alpha_1, \alpha_3\}$ in the above picture (again, TODO) is a set of minimal elements for the root system of SU_3, since α_2 can be written as a sum of the others.

(2) For U_n and $C=\{x_1>x_2>\cdots>x_n\}$, $\Phi^+=\{e_i-e_j \mid i<j\}$ and $\Phi^-=-\Phi^+$ (e.g. by taking $\lambda=(n,n-1,n-2,\dots,1)\in C$), and $B=\{e_i-e_{i+1}\}$, since (for example) $e_i-e_{i+2}=(e_i-e_{i+1})+(e_{i+1}-e_{i+2})$. Notice that B is a vector-space basis for $X^*\otimes\mathbb{R}$; this is not a coincidence. ◂

One can spend quite a lot of time on the geometry of root systems.

An alternate characterization of the minimal elements is, as hinted above, a basis.

Definition 22.3. A basis B for a root system Φ is an $\mathbb{R}$-basis for $\mathbb{R}\Phi$ such that every $\alpha\in\Phi$ can be written as

$$\alpha=\sum_{\gamma\in B} n_\gamma\gamma,$$

where $n_\gamma\in\mathbb{Z}$, $n_\gamma\geq 0$ when $\alpha\in\Phi^+$, and $n_\gamma\leq 0$ when $\alpha\in\Phi^-$.

This basis ends up being convenient for computing things in.

Theorem 22.4. *The minimal elements form a basis.*

Proof sketch. Clearly, every element in Φ is decomposable as a $\mathbb{Z}$-linear combination of elements of a minimal set B (just break them down into sums, and keep going until only minimal elements remain), so B spans $\mathbb{R}\Phi$.

The next step is to address linear independence. If $\alpha,\beta\in B$, then $\langle\alpha,\beta\rangle\leq 0$ (after we've chosen an invariant inner product); this is because if $\langle\alpha,\beta\rangle>0$, then $s_\beta\alpha\in\Phi$, which ends up implying (via the root system axioms, or the representation theory of SU_2)[42] that $\alpha-\beta\in\Phi$. Without loss of generality, $\alpha-\beta\in\Phi^+$, but then, $\alpha=(\alpha-\beta)+\beta$ isn't minimal, so this is a contradiction. Thus, by geometry, this implies the elements of B are linearly independent, because they are all at obtuse angles to each other, and all lie on the same side of a hyperplane. ⊠

The full proof can be found in the textbook.

Returning to representation theory, recall that we're trying to prove that the irreducible representations of a compact, connected Lie group G are parameterized by W-orbits of $X^*(T)$. For example, each orbit for U_n of $X^*(T)=\mathbb{Z}^n$ under $W=S_n$ corresponds to one irreducible representation; that is, each choice of $\lambda_1\geq\cdots\geq\lambda_n$ for $\lambda_i\in\mathbb{Z}$.

The representation V_χ corresponding to a given $\chi\in X^*$ has the property that

$$V_\chi|_T=\left(\bigoplus_{\chi'\in W_\chi}\chi'\right)\oplus(\text{smaller }\psi\in X^*(T)),$$

for a meaning of "smaller" that will be clarified. Furthermore, we will be able to write down an explicit character formula.

In the case $G=\mathrm{SU}_2$, T consists of matrices of the form $\left(\begin{smallmatrix} z & 0\\ 0 & z^{-1}\end{smallmatrix}\right)$, so let

$$\chi_m:\begin{pmatrix} z & 0\\ 0 & z^{-1}\end{pmatrix}\longmapsto z^m.$$

Then, the irreducible components are given in Table 4. Here we can see that "smaller" means components for which the length in some inner product is strictly less than that of χ. We could have used the explicit character formula for this, but it illustrates the complexity of the general case, and in fact it looks pretty, but nobody really uses it.

	representation on G	restriction to T
V_{χ_0}	1-dim. rep. of SU_2	χ_0
V_{χ_1}	2-dim. rep. of SU_2	$\chi_1\oplus\chi_{-1}$
V_{χ_2}	3-dim. rep. of SU_2	$\chi_2\oplus\chi_0\oplus\chi_{-2}$
V_{χ_3}	4-dim. rep. of SU_2	$\chi_3\oplus\chi_1\oplus\chi_{-1}\oplus\chi_{-3}$

TABLE 4. Table of irreducible components of a representation of SU_2.

We'll dive into the proof of these statements next week. It's all a generalization of what we can see in the case of the unitary group, which we'll recall right here.

[42] A third valid way to prove this would be to classify all five examples, but that's kind of tedious.

Outline of the proof for U_n. The key point is the Weyl integration formula. Let $T = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix}$ as usual and f be a class function on U_n, so that

$$\int_{\mathrm{U}_n} f = \frac{1}{n!} \int_T f \cdot \prod_{i \neq j} \left(\frac{z_i}{z_j} - 1 \right) \mathrm{d}t$$

$$= \frac{1}{n!} \int_T f \overbrace{\left| \prod_{i<j} (z_i - z_j) \right|^2}^{D}. \tag{22.5}$$

This will end up generalizing in the following way:

$$\int_G f = \frac{1}{|W|} \int_T f \cdot \prod_{\alpha \in \Phi} (\alpha(t) - 1) \, \mathrm{d}t.$$

In this general case, we will have to choose a chamber and some positive roots to simplify this.

Returning to the specific case of U_n, (22.5) says something interesting about the eigenvalue distribution of a random unitary matrix; the formula is different for other classical groups, and thus the resulting eigenvalue distribution differs.

If V is an irreducible U_n-representation, then $\langle \chi_V, \chi_V \rangle = 1$, so

$$\int_{\mathrm{U}_n} |\chi|^2 = \frac{1}{n!} \int_T |\chi D|^2,$$

which brings us to Fourier analysis, or the representation theory of the torus. On T, χ is a $\mathbb{Z}$-linear combination of elements of $X^*(T)$, because V splits into irreducibles over T, so there are some $m_1, \dots, m_n \in \mathbb{Z}^{\geq 0}$ such that

$$\chi \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} = \sum a_{\mathbf{m}} z_1^{m_1} \cdots z_n^{m_n}.$$

Furthermore, these $a_{\mathbf{m}}$ are symmetric in $(m_1, \dots, m_n)$.[43] Additionally, D is skew-symmetric,[44] Thus,

$$D\chi = \sum b_{\mathbf{m}} z_1^{m_1} \cdots z_n^{m_n},$$

where

$$b_{\sigma(\mathbf{m})} = \operatorname{sign}(\sigma) b_{\mathbf{m}}.$$

Thus, if $b_{\mathbf{m}} \neq 0$, then all of the m_i must be distinct, as swapping the identical ones would cause it to change sign. Thus, if it's nonzero, then $\sum |b_{\mathbf{m}}^2| = n!$, since one must be ± 1, and then adding all its permutations means there can't be anything else.

Thus, there exists an $\mathbf{m}_0$ such that $b_{\sigma(\mathbf{m}_0)} = \operatorname{sign}(\sigma)$, and $b_{\mathbf{m}} = 0$ if $\mathbf{m} \neq S_n \cdot \mathbf{m}_0$ (that is, it's not in the orbit). Hence, the character formula pops out:

$$\chi_V = \frac{\pm \sum_{\sigma \in S_m} \operatorname{sign}(\sigma) z_{\sigma(1)}^{m_1} \cdots z_{\sigma(n)}^{m_n}}{\prod_{i<j} (z_i - z_j)}.$$

Notice how this follows only from $\langle \chi_V, \chi_V \rangle = 1$. This miracle will hold on a general G, with the twist that D is only skew-symmetric up to a cocycle.

Moreover, since the χ_V form a basis, then all of the $\mathbf{m} = (m_1, \dots, m_n)$ with $m_1 > \cdots > m_n$ must arise; if not, then Fourier analysis on the torus would point out that something must be orthogonal to all of them, and then the remaining assertions about χ and expressing it in terms of smaller representations comes from calculations similar to those on Homework 1.

Finally, we will have to compute the Weyl integration formula, which comes down to calculating the Jacobian.

[43] That is, if one swaps z_1 and z_2, the same character comes out in the end, because those elements are conjugate. This comes from the Weyl group in general, where we can swap only those elements which are symmetric under it.

[44] This turns out to not quite be true in the general case, but is close enough to work.

Today's lecture was given by Zhiwei Yun.

Spin groups are double covers of SO_n. We'll deduce their existence, discuss their root systems, and provide an explicit construction.

Notice that SO_n is not simply connected, e.g. $\mathrm{SO}_2 \cong S^1$, which of course isn't simply connected, and $\mathrm{SO}_3 \underset{2:1}{\twoheadleftarrow} \mathrm{SU}_2 \cong S^3$ and is also the group of unit quaternions, so $\pi_1(\mathrm{SO}_3) \cong \mathbb{Z}/2\mathbb{Z}$.

Theorem 23.1. *$\pi_1(\mathrm{SO}_3) \cong \mathbb{Z}/2\mathbb{Z}$ if $n \geq 3$.*

Proof. Embed $\mathrm{SO}_n \hookrightarrow \mathrm{SO}_{n+1}$ via $A \mapsto \begin{pmatrix} A & \\ & 1 \end{pmatrix}$. We want to show this induces an isomorphism on fundamental groups.

The action of SO_{n+1} on $S^n \subset \mathbb{R}^{n+1}$ is transitive, and the stabilizer of $\mathbf{x} = (0, \dots, 0, 1)$ is $\mathrm{Stab}(\mathbf{x}) = \mathrm{SO}_n$, embedded as above. Thus, there's a fibration $\mathrm{SO}_{n+1} \twoheadrightarrow S^n$ by $g \mapsto g \cdot x$ (the preimage of any point is isomorphic to SO_n), and so one gets a fibration

$$\mathrm{SO}_n \hookrightarrow \mathrm{SO}_{n+1} \twoheadrightarrow S^n$$

which is in fact a short exact sequence, and furthermore is locally trivial. But this means there's a long exact sequence

$$\cdots \longrightarrow \pi_2(S^n) \longrightarrow \pi_1(\mathrm{SO}_n) \xrightarrow{\varphi} \pi_1(\mathrm{SO}_{n+1}) \longrightarrow \pi_1(S^n) \longrightarrow \pi_0(\mathrm{SO}_n).$$

$\pi_0(\mathrm{SO}_n) = 0$ because SO_n is connected. This long exact sequence is attached to any fibration, which can be seen from any textbook in homotopy theory. We want to show that φ is an isomorphism, but $\pi_1(S^n)$ and $\pi_2(S^n)$ are both trivial when $n \geq 3$, so this implies $\varphi : \pi_1(\mathrm{SO}_n) \xrightarrow{\sim} \pi_1(\mathrm{SO}_{n+1})$. Thus, since $\pi_1(\mathrm{SO}_3) \cong \mathbb{Z}/2$, then $\pi_1(\mathrm{SO}_n) \cong \mathbb{Z}/2$ for all $n \geq 3$. ⊠

This means that there's a connected double cover of SO_n, and it's a universal cover, by the general theory of covering spaces. And from the general theory of Lie groups, this universal cover $\widetilde{\mathrm{SO}_n}$ carries a Lie group structure; this will be the spin group Spin_n, and so it abstractly exists.

We saw above that $\mathrm{Spin}_3 \cong \mathrm{SU}_2$. For Spin_4, recall that $\mathbb{H}$ acts on itself from the left and right, and $\mathrm{SU}(2) \subset \mathbb{H}$ is the set of unit quaternions. Thus, $\mathrm{SU}(2) \times \mathrm{SU}(2)$ acts from the left and right on $\mathbb{H}$, and preserves the norm, so there's a homomorphism $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. Since both have dimension 6, then this is a covering map, and since $\mathrm{SU}(2) \times \mathrm{SU}(2)$ is both connected and simply connected, then it's the universal cover. Thus, $\mathrm{Spin}_4 \cong \mathrm{SU}_2 \times \mathrm{SU}_2$.

These coincidences arise when n is small: in general, it's not possible to write Spin_n in terms of other groups, and so we'll give a more general construction.

The Root System. We can actually figure out the abstract structure of the root system for Spin_n. Recall that the maximal torus T for SO_n is given by 2×2 rotation matrices (and possibly a 1 at the end), so there are $m = \lfloor n/2 \rfloor$ blocks, giving $\mathrm{SO}(2)^m$. Then, the maximal torus for Spin_n is just the preimage of the covering map:

$$\begin{tikzcd} \widetilde{T} \arrow[r, hook] \arrow[d] & \mathrm{Spin}_n \arrow[d] \\ T \arrow[r, hook] & \mathrm{SO}_n \end{tikzcd}$$

This requires showing that $\widetilde{T}$ is connected, but that follows from general covering space theory: it induces a surjection on π_1: $\mathrm{SO}(2) \to \mathrm{SO}(n)$ is sent to $\mathbb{Z} \to \mathbb{Z}/2$.

Thus, the roots of Spin_n lie in $X^*(\widetilde{T}) = \mathrm{Hom}(\widetilde{T}, S^1)$, which is a free abelian group with rank equal to that of $\widetilde{T}$, and the roots of SO_n lie in $X^*(T) = \mathrm{Hom}(T, S^1)$. But whenever we have a map $\widetilde{T} \to T$, Hom induces a map in the other direction, $X^*(T) \to X^*(\widetilde{T})$, so the roots of Spin_n satisfy {roots of $\mathrm{Spin}_n \subset X^*(T) \subset X^*(\widetilde{T})$. In general, the roots come from the action of the torus on the Lie algebra, so the roots of a Lie group and its universal cover are canonically identified.

Let $\varepsilon_i : T \to \mathrm{SO}(2) = S^1$ be the projection to the i^{th} copy of the lattice, so that $X^*(T) = \mathbb{Z}\varepsilon_1 \oplus \cdots \oplus \mathbb{Z}\varepsilon_m$, and recall that the roots of SO_n are $\pm\varepsilon_i \pm \varepsilon_j$ for $i \neq j$ if $n = 2m$ and additionally $\pm\varepsilon_i$ if $n = 2m+1$.

It remains to describe $X^*(\widetilde{T})$, which is bigger than $X^*(T)$, but only slightly so. Since $\widetilde{T} \to T$ is a 2:1 cover, so the index $[X^*(\widetilde{T}) : X^*(T)] = 2$. Thus, the concrete picture (TODO) is that of a sublattice of index 2 inside the lattice

$X^*(\widetilde{T})$. There are two ways to get a lattice of index 2, but if you multiply anything in $X^*(\widetilde{T})$ by 2, it lands in $X^*(T)$. Thus,

$$\overbrace{\underbrace{X^*(T) \subset X^*(\widetilde{T})}_{\text{index } 2} \subset \frac{1}{2}X^*(T)}^{\text{index } 2^m},$$

akin to $\mathbb{Z}[1/2]^m$ and $\mathbb{Z}^m$. Thus, $(1/2)X^*(T)/X^*(T) \cong X^*(T) \otimes (\mathbb{Z}/2\mathbb{Z})$, which is an $\mathbf{F}_2$-vector space of rank m, and $X^*(\widetilde{T})/X^*(T)$ is a line in said $\mathbf{F}_2$-vector space, the span of some vector.

To describe this, we'll use the action of the Weyl group, which is the same for Spin_n and SO_n.[45] In particular, the line $X^*(T)/X^*(\widetilde{T})$ is stable under the action of W on that vector space.

But the action of W on this vector space surjects to S_2 acting on it, because $W \twoheadrightarrow S_n$ acts on $\mathbf{F}_2 \cdot \varepsilon_1 \oplus \cdots \oplus \mathbf{F}_2 \cdot \varepsilon_m$. But only one line is invariant under this action: that spanned by $(1, \dots, 1)$, since this is the only eigenvector. Thus, the lattice gets uniquely pinned down, as

$$\begin{aligned} X^*(\widetilde{T}) &= \mathrm{Span}_{\mathbb{Z}} \left\{ \varepsilon_1, \varepsilon_2, \dots, \varepsilon_m, \frac{1}{2}(\varepsilon_1 + \cdots + \varepsilon_n) \right\} \\ &= \left\{ \sum a_i \varepsilon_i \mid a_i \in \frac{1}{2}\mathbb{Z} \right\}. \end{aligned}$$

The roots are identical, but they live in this slightly bigger lattice.

Construction of the Spin Groups. We'll be able to provide an explicit construction of these groups, but it's not that straightforward; we'll need to introduce some auxiliary objects, which are in themselves very important.

Let (V, q) be a quadratic space over $\mathbb{R}$ (i.e., q is a nondegenerate quadratic form, though we don't need it to be positive definite yet).

Definition 23.2. The Clifford algebra of (V, q) is $\mathrm{C}\ell(V, q) = T(V)/\langle v \otimes v + q(v) \cdot 1 \mid v \in V \rangle$ (recall that $T(V)$ is the tensor algebra of V, $\bigoplus_{i=0}^{\infty} V^{\otimes n}$, and that the angle brackets denote the two-sided ideal generated by the elements inside it).

For example, $v_1(3 + v_1 v_2 + v_1 v_2 v_3) = 3v_1 + (-q(v_1)v_2) + (-q(v_1))v_2 v_3$, so multiplication of any vector with itself yields a scalar.

This seems like quite an artificial object, but we've seen it before, albeit as a degenerate example: $\Lambda^*(V) = T(V)/\langle v \otimes v \rangle = Cl(V, 0)$; then, the general $\mathrm{C}\ell(V, q)$ can be seen as a deformation of $\Lambda^*(V)$. In particular, one can prove that they have the same dimension: $\dim(\mathrm{C}\ell(V, q)) = 2^n$ if $\dim(V) = n$.

If $e_1, e_2, \dots, e_n$ is an orthonormal basis for (V, q) (i.e. $q(e_i) = \pm 1$), then $\mathrm{C}\ell(V, q)$ has as a basis

$$\{ e_{i_1} e_{i_2} \cdots e_{i_k} \mid 1 \le i_1 \le i_2 \le \cdots \le i_k \le n \}.$$

Reordering only changes the value up to sign, because applying the relation to $e_i + e_j$ gives $e_i e_j + e_j e_i = 0$.

Letting $C := \mathrm{C}\ell(V, q)$, there's an obvious grading $C = C^0 \oplus C^1$ into even-degree and odd-degree parts, respectively, because the imposed relation has purely even degree. Thus, C^0 is a subalgebra, and acts on the odd part: C^1 is a C^0-module, and since $V \hookrightarrow T(V)$ still survives under the quotient (look at the explicit basis), then $V \subset C^1$ also.

We'll also need an involution on the Clifford algebra, called transposition, $x \mapsto x^{\mathrm{T}}$, which sends

$$e_{i_1} \cdots e_{i_k} \longmapsto e_{i_k} e_{i_{k-1}} \cdots e_{i_1}.$$

Like matrix transposition, this satisfies $(xy)^{\mathrm{T}} = y^{\mathrm{T}} \cdot x^{\mathrm{T}}$.[46]

Since $C = C^0 \oplus C^1$, then there must be an order-2 operator ε such that this is its eigenspace decomposition: $\varepsilon \cdot C^0 = 1$, and $e \cdot C^1 = -1$ (i.e. that's what ε restricts to on each part), and it's an algebra automorphism. Thus, we can define a "norm" $N(x) := x \cdot \varepsilon(x^{\mathrm{T}}) \in C$.

Exercise 23.3. Piece these things out in the case $\dim(V) = 2$, and where q is the standard quadratic form. In this case, $\mathrm{C}\ell(V, q) \cong \mathbb{H}$, and N becomes the usual norm.

[45] This technically needs to be checked, but follows from the definition: take the preimage of $N(T)$ under the covering map. This ends up holding true for any covering of Lie groups.

[46] This can actually be defined using the universal property of the Clifford algebra: it is the unique operator that restricts to the identity on V, and satisfies $(xy)^{\mathrm{T}} = y^{\mathrm{T}} x^{\mathrm{T}}$.

Definition 23.4. $\mathrm{Spin}(V,q) = \{x \in C^{0,\times} \mid xvx^{-1} \in V$ for every $v \in V$, and $N(x) = 1\}$. Here, $C^{0,\times}$ is the set of invertible elements in C^0 (though if $N(x) = 1$, x turns out to automatically be invertible).

Now, we have a spin group associated to any quadratic form q.

Theorem 23.5. *If q is positive definite, then $\mathrm{Spin}(V,q) \to \mathrm{SO}(V,q)$ is the universal cover.*

This map exists because anything in $\mathrm{Spin}(V,q)$ acts on V by conjugation within $C\ell(V,q)$, and this action preserves the quadratic form, so it ends up mapping to something in the orthogonal group: for any $x \in V \setminus 0$, viewed as in C^1, the action$v \mapsto (-x)vx^{-1}$ turns out to be invertible. Thus, Spin_n falls out for the standard case.

24. The Weyl Integration Formula: 5/28/14

"When I was a graduate, most of my classes had about three students. The goal was to avoid being the single remaining person in a class. It was very delicate."

Today, all integrals will be taken with respect to the Haar measure with total mass 1.

Recall that for U_n, the Weyl integration formula was

$$\int_{\mathrm{U}_n} f = \frac{1}{n!}\int_T f \prod_{i<j} |z_i - z_j|^2, \tag{24.1}$$

where T is the maximal torus of U_n (i.e. the diagonal matrices with entries in S^1), and f is a class function.

Proposition 24.2. *In general, if G is a compact Lie group and $T \subset G$, $\Phi \subset X^*(T)$, and W are defined as normal, then for a class function f on G,*

$$\int_G f(g) = \frac{1}{|W|}\int_T f(t) \prod_{\alpha\in\Phi} (\alpha(t) - 1). \tag{24.3}$$

When $G = \mathrm{U}_n$, this reduces to (24.1) in not very many steps.

The formula (24.1) means in words that if one chooses a $g \in \mathrm{U}_n$ according to the Haar measure, the probability distributions of the eigenvalues $z_1, \dots, z_n$ is

$$\frac{1}{n!}\prod_{i<j} |z_i - z_j|^2 \, \mathrm{d}z_1 \cdots \mathrm{d}z_n.$$

Thus, the eigenvalues tend to repel each other (which is still mostly true for more general G).

Proof of Proposition 24.2. This proof won't be that hard, and follows from the use of the Jacobian and a change of variables $G/T \times T \to G$, sending $(x,t) \mapsto xtx^{-1}$. This is more general than we need, though: we only care about the class function case, which simplifies a bit further.

The proof leans on the following fact.

Proposition 24.4. *Suppose M and N are compact, oriented manifolds of the same dimension d and $f : M \to N$ is smooth. Then, there exists an $m \in \mathbb{Z}$ such that for any volume form ω on M, $\int_M f^*\omega = m \int_N \omega$.*

This m is called the *degree* of f, and is equal to the number of preimages of a generic point in N. That is, for $x \in N$, let $y_1, \dots, y_n$ be its preimages in M, and count each one with a $+1$ if orientation is preserved or -1 if it isn't; this weighted sum is equal to m almost everywhere (according to the Lesbegue measure in each coordinate chart).

The example to keep in mind is $e^{i\theta} \mapsto e^{in\theta}$, $S^1 \to S^1$. The proof of this intermediate proposition is easy in the case of x covered by f (as if it were a covering map), and then the rest follows by deforming ω.

We'll use this for $G/T \times T \to G$. Let ω and μ be G-invariant (both left and right) volume form ons on G and G/T, respectively (i.e. differential forms of top degree), and ν be a T-invariant volume form on T. These will be useful because the Haar measure was constructed by integrating an invariant volume form.

We've discussed why ω and ν ought to exist: it can be first fixed at the identity as an element of $\Lambda^{\dim G}(\mathrm{Lie}\, G)^*$, and then define it elsewhere using G-invariance. But this is a bit of a problem on G/T, since there are multiple ways to left-translate, so to construct μ, one needs to show that T acts trivially on $\Lambda^{\dim(G/T)}(\mathrm{Lie}(G)/\mathrm{Lie}(T))$, which is equivalent to saying that the determinant of T acting on $\mathrm{Lie}(G)/\mathrm{Lie}(T)$ is 1. But this determinant is precisely

$$\det = \prod_{\alpha\in\Phi} \alpha,$$

and since the roots come in $\{x, x^{-1}\}$-pairs, then this works. It's hardly a triviality, but it's not all that important.

We can and should choose ω, μ, and ν to be compatible, so that

$$\int_G f(g)\,\mathrm{d}\omega = \int_{G/T}\int_T f\,\mathrm{d}\mu\,\mathrm{d}\nu,$$

or equivalently, requiring that they're all compatible at e, which is carefully worked out in the textbook:

$$\begin{aligned}
\omega|_e &\in \Lambda^{\dim(G)}(\mathrm{Lie}(G)^*),\\
\mu|_e &\in \Lambda^{\dim(G)-\dim(T)}(\mathrm{Lie}(G)/\mathrm{Lie}(T))^*,\\
\nu|_e &\in \Lambda^{\dim(T)}(\mathrm{Lie}(T)^*).
\end{aligned}$$

Recall that if W is an m-dimensional subspace of the n-dimensional vector space V, there's a canonical isomorphism $\Lambda^n V \cong (\Lambda^n W)\otimes(\Lambda^{n-m}(V/W))$; then, the goal is to make them compatible with that isomorphism.

Now, ω defines the Haar measure on G: $\int_G f = \int_G f\,\mathrm{d}\omega$. Let $\pi : G/T \times T \to G$ send $(x,t)\to xtx^{-1}$; this will end up being surjective, so that everything in G is conjugate to something in T, but we haven't shown that yet. Let $m = \deg(\pi)$ (as in the aforementioned proposition), so that

$$\begin{aligned}
\int_G f\,\mathrm{d}\omega &= \frac{1}{m}\int_{G/T\times T} \pi^*(f\omega)\\
&= \frac{1}{m}\int_{x\in G/T,t\in T} f(xtx^{-1})\pi^*\omega\\
&= \frac{1}{m}\int_{G/T\times T} f(t)J(x,t)\mu\wedge\nu,
\end{aligned}$$

where J is the Jacobian, defined as $\pi^*\omega = J(x,t)\mu\wedge\nu$, specifically evaluated at the point (x,t). Thus, the next step is to compute the Jacobian.

Because μ is G-invariant on G/T and ω on G is both left and right G-invariant (see the discussion on Haar measure for a reason why), then $J(x,t)$ is independent of x! This is because if $x,x'\in G/T$, then they can be translated to each other, but volume forms don't change, so this is unaffected. Thus, call the Jacobian simply $J(t)$. Thus, the formula now looks like this:

$$\begin{aligned}
\int_G f &= \frac{1}{m}\int_{G/T\times T} f(t)J(t)\mu\wedge\nu\\
&= \left(\frac{1}{m}\int_T f(t)J(t)\nu\right)\mathrm{Vol}(G/T).
\end{aligned}$$

Next, normalize (by scaling) ω and ν such that $\int_G\omega = 1$ and $\int_T\nu = 1$. Thus, by the compatibility condition, $\mathrm{Vol}(G/T) = 1$, so

$$\int_G f = \frac{1}{m}\int_T f(t)J(t),$$

where the right-hand side is also Haar measure.

There are two ingredients left now: J and m, and the computation of J is probably the most important part of the entire proof.

Since J is independent of x, we'll compute it at $x=1$, so at (eT,t). Then, explicitly linearize: let $\mathfrak{u}\subseteq\mathrm{Lie}(G)$ be the orthogonal complement to $\mathrm{Lie}(T)$, so $\mathfrak{u}\xrightarrow{\sim}\mathrm{Lie}(G)/\mathrm{Lie}(T)$.

Locally, near $(eT,t)\in G/T\times T$, this map looks like this for $X\in\mathfrak{u}$ and $Y\in\mathrm{Lie}(T)$:

$$\begin{aligned}
(e^X T, te^Y) &\longmapsto e^X te^Y e^{-X}\\
&= t(t^{-1}e^X t)e^Y e^{-X}\\
&= t\left(e^{\mathrm{Ad}(t^{-1})X}\right)e^Y e^{-X}\\
&\approx te^{\mathrm{Ad}(t^{-1})X+Y-X}
\end{aligned}$$

to first-order, since multiplication acts like addition. In effect, Ad is the Lie-algebraic version of conjugation.

In other words, the linearization of π at (eT, t) looks like

$$(X \in \mathfrak{u}, Y \in \mathrm{Lie}(T)) \longmapsto (\mathrm{Ad}(t^{-1}) - 1)X + Y \in \mathrm{Lie}(G) = \mathfrak{u} \oplus \mathrm{Lie}(T).$$

Now, we only have to compute the determinant: $J(t)$ is the determinant of $\mathrm{Ad}(t^{-1})$ acting on $\mathfrak{u}$ (since this map is the identity on Y). But the eigenvalues of $\mathrm{Ad}(t)$ on $\mathfrak{u} = \mathrm{Lie}(G)/\mathrm{Lie}(T)$ are the roots $\{\alpha(t)\}_{\alpha\in\Phi}$, just by definition, so since the roots come in inverse pairs,

$$J(t) = \prod_{\alpha\in\Phi}(\alpha(t)^{-1} - 1) = \prod_{\alpha\in\Phi}(\alpha(t) - 1).$$

This is the core of the computation. It says that if some root has $\alpha(t)$ near 1 (e.g. close eigenvalues in U_n), conjugation doesn't move it very much.

Now, the formula is nearly complete: it still has the degree, though.

$$\int_G f = \frac{1}{m}\int_T f(t)\prod_{\alpha\in\Phi}(\alpha(t) - 1)\,\mathrm{d}t.$$

In this formula, the Jacobian is positive because the pairs are $\alpha(t)$ and $\alpha(t)^{-1} = \overline{\alpha(t)}$, so the product is over a number of squared terms. Thus, π is positively oriented when $J(t)$ is nonzero (where the orientations were specified with the differential forms that were chosen).

Now, choose a $t_0 \in T$ generic in that $\alpha(t_0) \neq 1$ for all $\alpha \in \Phi$ and $\overline{\langle t\rangle_0} = T$ (e.g. in U_n, one requires there to be no repeated eigenvalues). Then, we'll figure out what m is by counting its preimages. If $(xT, t) \overset{\pi}{\mapsto} t_0$, then $xtx^{-1} = t_0$, so $x^{-1}t_0x = t$, and thus $x^{-1}\overline{\langle t_0\rangle}x \subset T$. Thus, $x^{-1}Tx \subset T$ so $x \in N(T)$. Conversely, each element of $N(T)$ creates a preimage $(x, x^{-1}t_0x)$ of t_0, and all of these are distinct (which is important to check). Thus, $m = |W|$. It remains to check that the set of these generic points has positive measure, so that it intersects the full-measure set (it does have full measure, of course, but it's not super relevant). ⊠

In U_n-land, this means that if D_1 and D_2 are conjugate diagonal matrices, then D_1 permutes the elements of D_2.

25. Skew-Symmetry and the Weyl Integration Formula: 5/30/14

"It's not fun, is it… you look sad. I was about to say 'it'll get better,' but maybe not."

Last time, we stated the Weyl integration formula, (24.2), for a compact connected Lie group G and class functions on it, and we showed that $G/T \times T \xrightarrow{\pi} G$ has degree $|W| > 0$, and thus π is surjective. Thus, every element of G is conjugate to an element of T. This is stronger than the result we had before, which was only that the maximal tori are conjugate to each other. In concrete cases, it boils down to the Spectral Theorem.

Corollary 25.1. *If S is a torus in G, then its centralizer $Z(S)$ is connected.*

Proof. Let $g \in Z(S)$ and $S' = \overline{\langle S, g\rangle}$; the latter is an abelian Lie group. Let $g' \in S'$ be generic, so that $\overline{\langle g'\rangle} = S'$.[47] Then, g' is conjugate to an element of T, so after conjugation, we can assume $S, g \subseteq T$, and therefore g can be joined to e inside $Z(S)$, because $T \subseteq Z(S)$ and g can be joined to e within T. ⊠

This corollary is very important in theory, but is a bit abstract to easily appreciate.

Corollary 25.2. $\exp : \mathrm{Lie}(G) \to G$ *is surjective.*

Proof. This is true for T, but then every element of G is conjugate to one in T. Alternatively, one could put a metric on G and think about minimal length paths. ⊠

Returning to representation theory, we want to write

$$\prod_{\alpha\in\Phi}(\alpha(t) - 1) = |D(t)|^2$$

for some D, which will make computations in the L^2-norm easier. Since the α come in conjugate pairs,this is possible, but it will require breaking symmetry in some way.

[47]If S' is disconnected, then let S'_e be the connected component of the identity, so that $S'/S'_e = \langle g\rangle$; thus, such a generic g' exists.

Let C be a chamber on $\mathrm{Lie}(T)^*$ (i.e. a connected component of $\mathrm{Lie}(T)-\bigcup\alpha^\perp$). This chamber gives a decomposition $\Phi=\Phi^+\cup\Phi^-$, wher $\Phi^+=\{\alpha\mid\langle\lambda,\alpha\rangle>0\}$ for some (or any) $\lambda\in C$, and $\Phi^-=\{\alpha\mid\langle\lambda,\alpha\rangle<0\}$. The minimal elements of Φ^+ (those that aren't the sums of others) give a basis for Φ; recall the picture from the case $G=\mathrm{SU}_3$.

Thus, we can write

$$\begin{aligned}\prod_{\alpha\in\Phi}(\alpha(t)-1)&=\prod_{\alpha\in\Phi^+}(\alpha^{-1}-1)\prod_{\alpha\in\Phi^-}(\alpha^{-1}-1)\\&=\left|\prod_{\alpha\in\Phi^+}(\alpha^{-1}-1)\right|^2\\&=\left|\prod_{\alpha\in\Phi^+}(1-\alpha^{-1})\right|^2.\end{aligned}$$

Call this quantity $D(t)$, so that the Weyl integration formula becomes

$$\int_G F=\frac{1}{|W|}\int_T|D(t)|^2F.$$

This is non-obvious already for constant functions: lots of things cancel.

Notice that for U_n, the formula we got is

$$D(t)=\prod_{i<j}\left(1-\frac{z_i}{z_j}\right),$$

rather than $\prod_{i<j}(z_i-z_j)$. This is annoying, because it's not strictly skew-symmetric, even though it comes from the more general formula. It's almost skew-symmetric, though… In this context, "skew-symmetric" should be generalized from S_n to W, so a skew-symmetric function f on G is such that whenever $t\in T$ and $w\in W$, $f(wt)$ is the determinant of w, acting on X^*, multiplied by $f(t)$. So we end up computing the W-action on D.

For $w\in W$,

$$wD=\prod_{\alpha\in\Phi^+}(1-w\alpha^{-1}(t)).$$

Some of the $w\alpha^{-1}(t)\in\Phi^-$, so it's not just a matter of rearranging signs. A bit more elegantly, rewrite it as

$$D=\prod_{\alpha\in\Phi^+}\left(\sqrt{\alpha}-\frac{1}{\sqrt{\alpha}}\right)\frac{1}{\sqrt{\alpha}}=\rho^{-1}\prod_{\alpha\in\Phi^+}\left(\sqrt{\alpha}-\frac{1}{\sqrt{\alpha}}\right),$$

where $\rho=\prod_{\alpha\in\Phi^+}\sqrt{\alpha}$ (or additively, it's $(1/2)\sum\alpha$, in $\mathrm{Lie}(T)^*$). This is the connection we need: it shows up across representation theory, as a spin structure somewhere. But what do $\sqrt{\alpha}$ and ρ mean? They make sense on a suitable covering of T, in the usual sense of square roots of holomorphic functions.

Under the action of w, there is some number of roots whose sign is flipped; call this number $\mathrm{sign}(w)$, so that

$$\begin{aligned}wD&=(w\rho)^{-1}\prod_{\alpha\in\Phi^+}\left(\sqrt{w\alpha}-\frac{1}{\sqrt{w\alpha}}\right)\\&=(w\rho)^{-1}\prod_{\alpha\in\Phi^+}\left(\sqrt{\alpha}-\frac{1}{\sqrt{\alpha}}\right)(-1)^{\mathrm{sign}(w)}\\&=\frac{\rho}{w\rho}\cdot D\cdot\mathrm{sign}(w).\end{aligned}$$

Proposition 25.3. $\mathrm{sign}(w)=\det(w\text{ on }X^*)$.

Thus, this would be skew-symmetric except for the $\rho/w\rho$ term. There's another fact about ρ which we'll want to have around for later.

Proposition 25.4. *If B is the basis associated to a chamber C, then $\langle\rho,\alpha^\vee\rangle=1$ for every $\alpha\in B$.*

For example, with U_n and the chamber given before, i.e. $C = \{a_1 > a_2 > \cdots a_n\}$ and basis $\{e_i - e_j \mid i < j\}$, then

$$\begin{aligned}\rho &= \frac{1}{2}\big((e_1-e_2)+(e_1-e_3)+\cdots+(e_1-e_n)+(e_2-e_3)+\cdots+(e_{n-1}-e_n)\big)\\ &= \frac{1}{2}((n-1)e_1+(n-3)e_2+\cdots+(1-n)e_n),\end{aligned}$$

so in standard coordinates,

$$\rho \longleftrightarrow \frac{1}{2}(n-1, n-3, n-5, \ldots, 1-n).$$

Then, in the basis B, $\langle \rho, e_i - e_{i+1}\rangle = 1$, since it's the difference of two adjacent entries.

Proof of Proposition 25.4. Recall that $\rho = (1/2)\sum_{\beta\in\Phi^+}\beta$, and that for $\alpha \in B$ and $\beta \in \Phi^+$, $s_\alpha\beta \in \Phi^+$ unless $\beta = \alpha$ (in which case, of course, $s_\alpha\alpha = -\alpha$). This is because everything in Φ^+ can be written as a linear combination of things in B, so the α term s negative, but everything else is still positive.

The point is, though, $s_\alpha\rho = \rho - \alpha$, so $\langle \rho, \alpha^\vee\rangle = 1$. ⊠

Proof sketch of Proposition 25.3. This also shows that if $w = s_\alpha$, then $\mathrm{sign}(w) = \det(s_\alpha)$. Then, $\langle s_\alpha : \alpha \in B\rangle$ generates W, which is the rough idea.

There's no good way around this ρ; it's a fact of life when doing representation theory of compact groups. Apparently it corresponds to taking the square root of the spinor bundle (?) on the flag variety, which in the case of Riemann surfaces is a spin structure.

So some clumsiness is inevitable in the general case, unlike for the unitary group. Thus, we'll declare a modified action $*$ of W on $X^*(T)$, defined by taking $w * \chi = w(\chi+\rho) - \rho$ on $X^*\otimes\mathbb{R}$, i.e. w acts in the same way, but the origin has been shifted to $-\rho$, so that $-\rho$ is fixed by this action.

For $\mathrm{SU}(2)$, $T = \left(\begin{smallmatrix} z & 0\\ 0 & z^{-1}\end{smallmatrix}\right)$ and $X^*(T) \cong \mathbb{Z}$, via $m \mapsto (z \mapsto z^m)$, with $\Phi = \{2,-2\}$: $\Phi^+ = \{2\}$ and $\Phi^- = \{-2\}$. This means $\rho = 1$, so the old action of W is a negation, but the new action is reflection about the point -1.

By modifying in this way, D is literally skew-symmetric: $w * D = \mathrm{sign}(w)D$. D is a linear combination of characters, as

$$D = \prod(1-\alpha^{-1}) \in \mathbb{Z}[X^*] = \sum n_\chi \chi,$$

so $w*$ applies this new action on all of them; thus,

$$w * D = \sum_\alpha n_\alpha(w * \chi).$$

This is nice, because $w*$ was defined to make it come through.

For example, in $\mathrm{SU}(2)$ again, $D(t) = 1 - a^{-1}(t)$, where $\{\alpha\} = \Phi^+$. This doesn't look symmetric, until one recenters at ρ.

Next time, we'll take the proof from U_n and adapt it to general G, which will be much easier now that it is skew-symmetric.

26. Irreducible Representations: 6/2/14

It's now time for one of the big theorems of the course.

Theorem 26.1. *Let G be a compact, connected Lie group, and denote $T \subset G$ and $\Phi \subset X^*$ as usual. Let $C \subset X^*\otimes\mathbb{R}$ be a chamber inducing $\Phi = \Phi^+ \cup \Phi^-$ and $\rho = (1/2)\sum_{\Phi^+}\alpha$. Then:*

1. *The irreducible representations of G are parameterized by $X^*(T)/W$, i.e. Weyl orbits on X^*. Each Weyl orbit on X^* has a unique representative in $\overline{C}$ (the closure of the chamber).*
2. *Recall that for $w \in W$ and $\chi \in X^*\otimes\mathbb{R}$, $w * \chi = w(\chi+\rho) - \rho$ and $\mathrm{sign}(w) = \det(w \text{ on } X^*) \in \{\pm 1\}$ (since it preserves a quadratic form). Then, for any $\chi \in X^* \cap \overline{C}$, let V_χ be the corresponding irreducible representation, so that the character of V_χ on T is*

$$\mathrm{character}_T(V_\chi) = \frac{\sum_{w\in W} w * \chi\,\mathrm{sign}(w)}{\sum_{w\in W} w * 1\,\mathrm{sign}(w)} = \frac{\sum_{w\in W} w(\chi+\rho)\mathrm{sign}(w)}{\sum_{w\in W} w\rho}.$$

1, *sometimes denoted* **1**, *is the trivial character for* T. *Moreover,*

$$\sum_{w\in W} w * \mathbf{1}\operatorname{sign}(w) = D = \prod_{\alpha\in\Phi^+}(1-\alpha^{-1}).$$

(3) *There is a dimension formula polynomial in* χ*:*

$$\dim V_\chi = \frac{\prod_{\alpha\in\Phi^+}\langle \chi+\rho,\alpha\rangle}{\prod_{\alpha\in\Phi^+}\langle\rho,\alpha\rangle}.$$

(4) *Finally, when one writes*

$$\operatorname{char}(V_\chi) = \sum_{\alpha\in X^*(T)} m_\alpha \alpha,$$

so that m_α *denotes the multiplciity of irreducibles, then* $m_\alpha = 0$ *if* $|\alpha| > |\chi|$,[48] *and if* $|\alpha| = |\chi|$, *then* $m_\alpha \le 1$, *with equality only if* $\alpha \in W\cdot\chi$.

Example 26.2. On U_n, recall that

- $X^*(T) \cong \mathbb{Z}^n$,
- $\Phi = \{e_i - e_j\}$,
- One chamber is $C \subset \mathbb{R}^n$ is given by $\{a_1 > \cdots > a_n\}$, so its closure is $\overline{C} = \{a_1 \ge a_2 \ge \cdots \ge a_n\}$.
- Then, $\Phi^+ = \{e_i - e_j \mid i < j\}$,
- $W = S^n$, and
- $\rho = (1/2)(n-1, n-3, \ldots, 1-n)$.

Then, the W-orbits of $(m_1,\ldots,m_n)\in\mathbb{Z}^n$ are up to permutation, so there's a unique representation parameterized by $m_1 \ge \cdots \ge m_n$.

For part (2), the character formula for U_n becomes

$$\chi_{V_{m_1,\ldots,m_n}}\begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n\end{pmatrix} = \frac{\sum_{\sigma\in S_n} z_{\sigma(1)}^{m_1+(n-1)/2}\cdots z_{\sigma(n)}^{m_n-(n-1)/2}\operatorname{sign}(\sigma)}{\sum_{\sigma\in S_n} z_{\sigma(1)}^{(n-1)/2}\cdots z_{\sigma(n)}^{(n-1)/2}\operatorname{sign}(\sigma)}$$
$$= \frac{\det(z_i^{m'_j})}{\det(z_i^{\mu'_j})},$$

where $m'_1 = m_1 + (n-1)/2$, $m'_2 = m_2 + (n-3)/2$, and so on, and $\mu'_1 = (n-1)/2$, $\mu'_2 = (n-3)/2$, and so forth. This looks a lot like the character formula we derived in the first week; there are minor differences that correspond to some rearrangement.

The dimension formula from part 3 is that if $\chi * \rho = (m'_1,\ldots,m'_n)$, then

$$\dim V_{(m_1,\ldots,m_n)} = \frac{\prod_{i<j}(m'_i - m'_j)}{\prod_{i<j}(j-i)}$$
$$= \frac{\prod_{i<j}(m_i - m_j + j - i)}{\prod_{i<j}(j-i)}.$$

◀

Example 26.3. $G = \mathrm{SO}(5)$ is another example in which things are small enough to be made explicit. Recall the dramatis personæ that we have derived for SO(5):

- $$T = \begin{pmatrix} \operatorname{rot}\theta_1 & 0 & 0\\ 0 & \operatorname{rot}\theta_2 & 0\\ 0 & 0 & 1\end{pmatrix}.$$
- $X^*(T) = \mathbb{Z}^2$.
- $\Phi = \{\pm e_1 \pm e_2, \pm e_1, \pm e_2\}$.
- $C = \{a_1 > a_2 > 0\}$, so $\overline{C} = \{a_1 \ge a_2 \ge 0\}$.

[48] These absolute values require fixing a G-invariant inner product, which we can do; this is usually done more algebraically.

- Thus, $\Phi^+ = \{e_1+e_2, e_1, e_1-e_2, e_2\}$, and
- $W = S_2 \cdot \{\pm 1\}^2$ (the roots can be flipped and two can be switched. This is the symmetry group of the picture of the root system).
- $\rho = 3e_1/2 + e_2/2 = (3/2, 1/2)$.

The W-orbits on X^* are therefore $\{a \geq b \geq 0 \mid a, b \in \mathbb{Z}\}$ (since (a,b) are up to orbit and up to sign). Denote the associated representation $V_{a,b}$, with character $\chi = (a,b)$. This means that $\chi + \rho = (a+3/2, b+1/2)$, and the dimension formula is

$$\dim V_{a,b} = \frac{(a+b+2)(a-b+1)(a+3/2)(b+1/2)}{3/2}.$$

The denominator is the same as the numerator, but with $a = b = 0$. It's not obvious, but this formula does end up being a positive integer.

Let $x = e^{i\theta_1}$ and $y = e^{i\theta_2}$ for concision, so

$$\rho D = x^{3/2}y^{1/2}\left(1-\frac{1}{xy}\right)\left(1-\frac{y}{x}\right)\left(1-\frac{1}{x}\right)\left(1-\frac{x}{y}\right),$$

and

$$\sum w\rho \cdot \operatorname{sign}(w) = x^{3/2}y^{1/2} - y^{3/2}x^{1/2} - x^{-3/2}y^{1/2} - x^{3/2}y^{-1/2} + \cdots$$

and the two sides end up being equal, so

$$\sum_{w \in W} w * \mathbf{1} \operatorname{sign}(w) = D$$

after all. This means that the general character formula is

$$\operatorname{char}(V_{a,b}) = \frac{x^{a+3/2}y^{b+1/2} + (\text{the other seven } W\text{-orbits})}{x^{3/2}y^{1/2} + (\text{the other seven } W\text{-orbits})}.$$

This works for decomposing any representation into irreducible ones; for example, let V be the representation of SO_5 on $\Lambda^2\mathbb{C}^5$.

First look at V restricted to the maximal torus; in $\mathrm{GL}_n(\mathbb{C})$,

$$\begin{pmatrix} \operatorname{rot}\theta_1 & & \\ & \operatorname{rot}\theta_2 & \\ & & 1 \end{pmatrix} \sim \begin{pmatrix} x & & & & \\ & -x & & & \\ & & y & & \\ & & & -y & \\ & & & & 1 \end{pmatrix},$$

where $\sim$ denotes conjugacy; thus, those are its eigenvalues. Let $\{f_1, \dots, f_n\}$ be a basis for $\mathbb{C}^5$; then, the character of V on the torus is

$$\chi_{V|_T} = xy + \frac{x}{y} + \frac{y}{x} + \frac{1}{xy} + x + x^{-1} + y + y^{-1} + 2.$$

These correspond, respectively, to the $f_1 \wedge f_3$-term, then $f_1 \wedge f_4$, then $f_3 \wedge f_2$, and so on. There are ten of these terms, since $\Lambda^2\mathbb{C}^5$ is 10-dimensional.

This means that the α such that $m_\alpha \neq 0$ are $(1, \pm 1)$, $(-1, \pm 1)$, $(\pm 1, 0)$, $(0, \pm 1)$, and $(0,0)$, the last with multiplicity 2. V must contain $V_{(1,1)}$, since we know the largest weight that will appear; then,

$$\dim(V_{1,1}) = \frac{4 \cdot 1 \cdot 5/2 \cdot 3/2}{7/2} = 10 = \dim V,$$

so $V = V_{(1,1)}$ must be irreducible.

If we had instead found that $V_{(1,1)} \subsetneq V$, then we would have computed its character, subtracted it, and continued.

Here, $V_{(a,b)} \subseteq V_{(1,0)}^{\otimes(a-b)} \otimes V_{(1,1)}^{\otimes b}$, where $V_{(1,0)}$ is the standard representation. This is a nice result: any representation of SO(5) is contained in a tensor product of small representations. But the right-hand side is huge; one has to extract the left-hand side somehow. ◀

Proof of Theorem 26.1. For the rest of this lecture, we'll begin the proof. Some things will be skipped over for time's sake, but are the easier, computational parts.

Let V be irreducible, so that $\langle \chi_V, \chi_V \rangle = 1$, and start with the Weyl integration formula.

$$\int_G |\chi_V|^2 = \frac{1}{|W|} \int_T |\chi_V \cdot D|^2.$$

D is a way of taking a square root, in some sense.

On the torus T, χ_V must be a nonnegative linear combination of the irreducible representations:

$$\chi_V|_T = \sum_{\alpha \in X^*(T)} m_\alpha \alpha,$$

where $m_\alpha \geq 0$. Since χ is a class function, then $m_{w\cdot\alpha} = m_\alpha$ for any $w \in W$. We've also rigged $w*$ such that $w * D = \mathrm{sign}(w) D$ and

$$D = \sum_{\alpha \in X^*(T)} n_\alpha \alpha,$$

with $n_\alpha \in \mathbb{Z}$ (not always positive). $w\chi_V = \chi_V$, so it's symmetric, but D is skew-symmetric.

Finally, write

$$\chi_V D = \sum_{\alpha \in X^*(T)} k_\alpha \alpha.$$

Now, $k_\alpha \in \mathbb{Z}$, and this quantity is skew-symmetric, as $k_{w*\alpha} = k_\alpha \,\mathrm{sign}(w)$.

No α appearing in $\chi_V D$ (i.e. such that $k_\alpha \neq 0$) can be fixed by any nontrivial element of W, because W permutes the chambers simply transitively, so if α is fixed by a nontrivial $w \in W$, then it's perpendicular to some root β. This means that $\langle \chi_V D, \chi_V D \rangle = \sum k_\alpha^2 = |W|$ (since, as soon as one is in, all of the rest follow), but $\sum k_\alpha^2 \geq |W|$, with equality only when

$$\chi_V D = \pm \sum_{w \in W} \mathrm{sign}(w) w * \alpha.$$

Thus, this is exactly the situation that happens: every irreducible character is of the form

$$\chi_V|_T = \frac{\pm \sum_{w \in W} w * \alpha \,\mathrm{sign}(w)}{D},$$

for some $\alpha \in X^*(T)$, and where no nontrivial element of W fixes α.

It remains to show the following.

(1) We have to show that every α occurs.
(2) It'll be necessary to pin down the sign.
(3) Finally, we'll have to make the highest-weight calculation.

27. Irreducible Representations, part II: 6/4/14

Recall that we're trying to prove Theorem 26.1 from last time; we'll keep the same notation.

Continuation of the proof of Theorem 26.1. Last time, we showed that if V is irreducible, then

$$\chi_V = \pm \frac{\sum w * \psi \,\mathrm{sign}(w)}{D}, \tag{27.1}$$

where $(W, *)$ acts freely on ψ (i.e. the stabilizer is $\{e\}$), and that every free $(W,*)$-orbit[49] occurs in the formula (27.1) for some irreducible V, which follows because the characters span a dense subset of class functions: if not, then there's some $\psi_0 \in X^*(T)$ orthogonal to all of them, so

$$\langle \chi_V D, \sum_{w \in W} w * \psi_0 \,\mathrm{sign}(w) \rangle = 0,$$

so the whole orbit is diagonal. This is a skew-symmetric function on the terms, but the span of $\chi_V D$ for irreduicble V is dense in the skew-symmetric functions on $L^2(T)$ (i.e. those that transform under W in the same way D does; in the case of U_n, these are literally skew-symmetric). This ends up leading to a contradiction.

[49] A free orbit means that if $w * \chi = \chi$, then $w = e$.

Specifically, since $\langle \chi_V \rangle$ is dense in $L^2(G)$, then it's sufficient to show that $C(G) \to L^2(T)$ given by $f \mapsto f|_T \cdot D$ has dense image in the skew-symmetric functions in $L^2(T)$. But this follows because (and there's an argument to be made here) the restriction $C(G) \twoheadrightarrow C(T)^W$ is surjective, i.e. any W-invariant continuous function on T extends to G.

This follows, albeit non-obviously, from the fact that every element can be conjugated into T, so defining this extension is easy (conjugate an arbitrary $g \in G$ into T, and then apply the function), but what's trickier is proving it's well-defined. This has to do with the conjugacy of maximal tori. The point is, every ψ has to occur, or there wouldn't be enough characters to span the space of class functions.

So now we know that characters are parameterized by free $(W,*)$-orbits on X^*, but the theorem dealt with the usual action. This is finessed by revealing the bijections $a : X^* \cap \overline{C} \to X^*/W$ sending $\chi \mapsto W\chi$, and in the other direction $b : X^* \cap \overline{C} \to X^*/(W,*)$, sending $\chi \mapsto W * \chi$. This allows the $*$-action to go away. For want of time, however, we'll only prove the surjectivity of b (the rest are similar): that every free $(W,*)$-orbit has a unique representative in $\overline{C}$.

For SU_2, $X^* = \mathbb{Z}$, and the $*$-action reflects about $\rho = -1$. The point is that every free orbit under $(W,*)$ (i.e. not -1) has a representation in $X^* \cap \overline{C} = \mathbb{Z}_{\geq 0}$, as does every orbit of W.

Pick a $\psi \in X^*$ such that its $(W,*)$-orbit is free, so there exists a w such that $w * \psi \in -\rho + C$. That is, since the Weyl group acts simply transitively on the set of chambers, so bring it to your favorite chamber. It can't be on a wall, because then it would be fixed by that reflection, so not free, so it must be in $-\rho + \overline{C}$. Here, $-\rho$ is involved somehow in passing between C and $\overline{C}$.

Let B be a basis for the root system associated to C, i.e. the minimal elements of Φ^+. Then, for every $\alpha \in B$, $\langle w * \psi, \alpha^\vee \rangle = \langle -\rho, \alpha^\vee \rangle + \langle \lambda, \alpha^\vee \rangle$ for some $\lambda \in C$. Since $\langle \lambda, \alpha^\vee \rangle > 0$ because $\alpha \in \Phi^+$, and we computed earlier that $\langle -\rho, \alpha^\vee \rangle = -1$, then $|ang * w * \psi, \alpha^\vee \geq 0$. In other words, there exists a w such that $\langle w * \psi, \alpha^\vee \rangle \geq 0$ for all $\alpha \in B$, but the set of λ such that $\langle \lambda, \alpha \rangle \geq 0$ is exactly $\overline{C}$, and in particular, $w * \psi \in \overline{C}$, so b is surjective.

a is trickier, because it's not free, but the proof can be done fairly easily by looking at the classical root systems.

Now, we know that the irreducible representations are parameterized by $\psi \in X^*/W$, with

$$\psi \in \overline{C} \longleftrightarrow \mathrm{char}(V_\psi) = \frac{\pm \sum \mathrm{sign}(w) w * \psi}{D}.$$

The next step is to modify this formula into the one we want.

Look at $V_\psi|_T$. It decomposes into a sum of irreducibles (on T):

$$V_\psi = \bigoplus_{\theta \in X^*} m_\theta \theta,$$

and the characters form a sum.

$$\mathrm{char}(V_\psi) = \sum_{\theta \in X^*} m_\theta \theta.$$

The $\theta \in X^*$ such that $m_\theta \neq 0$ are called the weights of V_ψ.

Let $R = \max_{\text{weights } \theta} \langle \theta, \rho \rangle$, and $\theta_1, \dots, \theta_n$ be the maximizing θ_i, i.e. $\left\langle \theta_j, \rho \right\rangle = R$. We'll keep track of what happens to these maximal weights in this formula; in particular,

$$((m_1\theta_1 + \cdots + m_n\theta_n) + (\text{lower-order terms})) \cdot D = \pm(\mathrm{sign}(w) w * \psi).$$

But D has inverses of positive roots, which are lower terms: if $\alpha \in \Phi^+$, then $\langle -\alpha, \rho \rangle < 0$. Thus, the left-hand side looks like

$$m_1\theta_1 + \cdots + m_n\theta_n + \text{lower-order terms},$$

and the right-hand side is simularly ψ plus lower-order terms (there is something to prove here). In particular, there is exactly one maximizing θ_j, namely $\theta_j = \psi$, and $m_\theta = 1$, with sign $+1$ (in the character formula). This is significant.

So the goal is to make $\overline{C}$ act like C: the claim is that $\langle \psi, \rho \rangle > \langle w * \psi, \rho \rangle$ for $w \neq e$, or equivalently,

$$\begin{aligned} &\langle \psi + \rho, \rho \rangle > \langle w(\psi + \rho), \rho \rangle \\ \iff &\langle \psi + \rho, \rho \rangle > \left\langle w + \psi, w^{-1}\rho \right\rangle \\ \iff &\left\langle w + \rho, \rho - w^{-1}\rho \right\rangle > 0. \end{aligned}$$

But $\psi + \rho \in C$ (since $\psi \in \overline{C}$, and translation by ρ moves it into the interior). Furthermore, ρ is a sum of positive roots and w^{-1} makes some of them negative, so $\rho = w^{-1}\rho$ is a sum of elements of Φ^+, but by definition, $\langle C, \Phi^+ \rangle > 0$.

Now, in this version of the formula, it's much clearer what actual representations look like. In particular, the trivial representation is V_1, with the trivial character $\psi = 1$. That is,

$$\begin{aligned} 1 &= \frac{\sum \operatorname{sign}(w) w * \mathbf{1}}{D} \\ \iff D &= \sum \operatorname{sign}(w) w * \mathbf{1}. \end{aligned}$$

This is not at all obvious combinatorially. We also have the final form of the character formula, which we proved last time.

Now let's talk about maximality; inside V_ψ, let $V_\psi[\alpha] = \{v \in V_\psi : tv = \alpha(t)v \text{ for } t \in T\}$, so that

$$V = \bigoplus_\alpha V_\psi[\alpha].$$

This corresponds to picking out a single irreducible component of V_ψ on the torus.

We showed that $\dim(V_\psi[\alpha]) = m_\alpha$, so $\dim(V_\psi[\psi]) = 1$. Additionally, if $\beta \in \Psi$, then $\mathfrak{g}_\beta V_\psi[\alpha] \subseteq V_\psi[\alpha+\beta]$ (where, you may recall, $\mathfrak{g}_\beta$ is the β-root space in $\mathfrak{g} = \operatorname{Lie}(G) \otimes \mathbb{C}$). The proof is identical to the proof that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

For $X \in \operatorname{Lie}(G)$, X acts on $v \in V_\psi$ by $X \cdot v = \frac{\mathrm{d}}{\mathrm{d}t} e^{tX} v$. Therefore, in particular, let $V_\psi[\psi] = \mathbb{C} \cdot x_\psi$, i.e. it's spanned by a single vector. Then, $\mathfrak{g}_\beta \cdot x_\psi = 0$ when $\beta \in \Phi^+$. (Remember, in the case $G = \mathrm{SU}_2$, we had (13.2); then, x_ψ is sort of like one of the endpoints.) x_ψ is sometimes called the highest weight vector, and ψ the highest weight. This holds in several senses; we have shown it extremizes the inner product with ρ.

Thus, $\psi * \beta$ is not a weight, since its inner product would be greater than $\langle \psi, \rho \rangle$, which is maximal.

Finally, if $m_\alpha > 0$, then I want to show that $|\alpha| \le |\psi|$, with equality iff $\alpha \in W\psi$. Here, we have to reason a little more with the Lie algebra: every weight α with $m_\alpha \ne 0$ is of the form

$$\alpha = \psi - \sum_{\beta \in \Phi^+} n_\beta \beta,$$

going down the same way we did with $m_\theta \theta$, ultimately because V_ψ is irreducible. Thus, $V_\psi = \operatorname{span}(X_1 \cdots X_m x_\psi)$ for $X_1, \dots, X_n \in \operatorname{Lie}(G)$. Without loss of generality, $X_i \in \mathfrak{g}_\beta$ for $\beta < 0$ or in $\operatorname{Lie}(T)_\mathbb{C}$ (which requires an argument: commute something positive over to $X_m x_\psi$, which kills it, and then there are some Lie brackets left over). But the point is, this is $V_\psi[\psi - \beta_1 - \cdots - \beta_m]$, so every weight can come from the top by going down.

With an invariant inner product, arrange the weights so that α maximizes $\langle \alpha, \alpha \rangle$. Then, replacing α by $w\alpha$ for $w \in W$, we can assume $\alpha \in \overline{C}$, and $\alpha = \psi - \sum_{\Phi^+} n_\beta \beta$, and $\psi \in \overline{C}$ too, and thus $\psi = \alpha + \sum_{\Phi^+} n_\beta \beta$. Thus,

$$\langle \psi, \psi \rangle = \langle \alpha, \alpha \rangle + \left\langle \sum n_\beta \beta, \sum n_\beta \beta \right\rangle + \text{cross terms}$$

(which are positive in a positive chamber), so $\sum n_\beta \beta = 0$, since $\langle \alpha, \alpha \rangle$ is maximal; thus, $\psi = \alpha$ again. ⊠

To actually construct representations, generalize the following argument: the n-dimensional representation of SU_2 consists of polynomials of degree n on $\mathbb{C}^2$, i.e. sections of a certain line bundle on $\mathbb{P}^1(\mathbb{C})$ (so not quite functions). Note that $\mathbb{P}^1(\mathbb{C}) \cong \mathrm{SU}_2/T$; in general, G/T has a structure as a complex manifold (for example, its real dimension is even), and representations of G can be constructed as sections of a line bundle on this manifold.