REPRESENTATIONS OF LIE GROUPS

PAUL D. NELSON

ABSTRACT. Notes for a semester course on the representation theory of Lie groups, focusing on compact groups and complex reductive groups, taught at ETH Zürich in Spring 2019.

CONTENTS

In mathematics and related fields, one often encounters vector spaces V, such as
- the vector space $V = L^2(X, \mu)$ consisting of functions $f : X \to \mathbb{C}$ on some nice topological space X that are square-integrable with respect to some measure μ, or
- the space of solutions ϕ to some linear differential equation $D\phi = 0$,

and so on. The inputs to the construction of V are sometimes invariant by some symmetry group G. In the first example, G might have a measure-preserving action on (X, μ). In the second, G might commute with the operator D (for instance, many of the basic equations of mathematical physics are invariant by something like the rotation group $\mathrm{SO}(3)$). In either case we obtain a linear action of G on V, i.e., a homomorphism $G \to \mathrm{GL}(V)$. The case that G is a Lie group is simultaneously one of the most interesting cases (because of the many examples) and accessible (because, e.g., connected Lie groups are in many respects simpler than finite groups, thanks to powerful tools from calculus and linear algebra). In this course, we still develop the basic theory of such actions. We emphasize the general theory of compact groups, their relation with complex reductive groups, the structure theory and representations of compact Lie groups, and conclude with a bit about representations of complex reductive groups.

Acknowledgments. We thank Subhajit Jana, Constantin Kogler, Kaj Bäuerle and many other participants of the course for corrections and comments.

1. Prerequisites

We recall here some of the basics from linear algebra, topology, differential geometry, Lie theory, and functional analysis that will be used throughout the course.

We won't assume any prior knowledge of representation theory.

1.1. **Conventions.** In this course,

$$\text{"topological space"} := \text{"separable Hausdorff topological space."}$$

The examples we consider will moreover be locally compact, but we will mention this hypothesis explicitly when it is used. Similarly

$$\text{"vector space"} := \text{"complex vector space"}$$

unless we specify otherwise.

1.2. Topological groups.

Definition 1. A *topological group* is a group G equipped with a topology such that the maps

$$G \times G \to G$$
$$(x, y) \mapsto xy$$

and

$$G \to G$$
$$x \mapsto x^{-1}$$

are continuous. Similarly, a *topological vector space* is a vector space equipped with a topology such that the maps

$$V \times V \to V$$
$$(x, y) \mapsto x + y$$

and

$$\mathbb{C} \times V \to V$$
$$(\lambda, x) \mapsto \lambda x$$

are continuous. We write in that case $\mathrm{GL}(V)$ for the group of continuous linear maps $\phi : V \to V$ that admit a continuous two-sided inverse.

For instance, any finite group (always equipped with the discrete topology) is a topological group, as is any Lie group; any finite-dimensional vector space is naturally a topological vector space, as is any Hilbert space or Banach space.

1.3. Lie theory.
We'll assume a large number of facts from basic Lie theory. For instance, a Lie group G has a Lie algebra $\mathfrak{g}$. A continuous homomorphism $G_1 \to G_2$ between two Lie groups is automatically smooth, or even analytic, and differentiates to a Lie algebra homomorphism $\mathfrak{g}_1 \to \mathfrak{g}_2$. For a finite-dimensional vector space V, the group $\mathrm{GL}(V)$ is a Lie group with Lie algebra $\mathrm{End}(V)$.

1.4. Haar measure.
We abbreviate "locally compact topological group" to "locally compact group," and similarly for "compact group."

Definition 2. Let G be a locally compact group. Recall that a *Radon measure dg* on G is a functional

$$C_c(G) \to \mathbb{C}$$
$$f \mapsto \int_G f \, dg = \int_{g \in G} f(g) \, dg$$

such that $f \geq 0 \implies \int_G f \, dg \geq 0$.[1] For each such f and each $h \in G$, we denote by $\lambda_h f$ and $\rho_h f$ the left and right translates of G, normalized so that $\lambda_{h_1 h_2} = \lambda_{h_1} \lambda_{h_2}$ and $\rho_{h_1 h_2} = \rho_{h_1} \rho_{h_2}$:

$$\lambda_h f(g) := f(h^{-1}g), \quad \rho_h f(g) := f(gh^{-1}).$$

Recall that dg is a *left Haar measure* if $\int_G f \, dg = \int_G \lambda_h f \, dg$ for all $h \in G$. The notion of a *right Haar measure* is defined analogously using ρ_h. We say that dg is a *Haar measure* if it is both a left Haar measure and a right Haar measure.

---

[1]The continuity and related conditions follow from the positivity, see, e.g., the Wikipedia entries on "Radon measure" and "Riesz-Kakutani representation theorem"

Theorem 3. *Let G be a locally compact group. Then left Haar measures exist, and any two are positive multiples of one another. Similarly for right Haar measures.*

Suppose moreover that G is compact. Then Haar measures exist, and assign finite volume to G.

In particular, there is a unique Haar measure dg on G such that $\mathrm{vol}(G, dg) = 1$; we call it the *probability Haar measure on G*. We will often denote integration with respect to the probability Haar simply by $\int_G f := \int_G f\, dg$, omitting the dg when it is clear by context.

For example, if G is finite, then the probability Haar dg is given by the normalized counting measure:

$$\int_G f\, dg = \frac{1}{|G|} \sum_{g \in G} f(g).$$

1.5. **Spectral theory for compact self-adjoint operators.** Let $V = (V, \langle, \rangle)$ be a Hilbert space and $T : V \to V$ a bounded operator. Recall that T is

- *self-adjoint* if $\langle Tv_1, v_2 \rangle = \langle v_1, Tv_2 \rangle$ for all $v_1, v_2 \in V$, and
- *compact* if it maps the unit ball to a precompact set, or equivalently, if each bounded sequence v_n in V has a subsequence v_{n_k} such that the sequence Tv_{n_k} converges.

A basic example is when T is diagonalized by an orthonormal basis $e_1, e_2, \ldots$ of V, thus $Te_j = \lambda_j e_j$ for some $\lambda_j \in \mathbb{C}$; then T is compact precisely when $\lambda_j \to 0$ as $j \to \infty$.

Theorem 4. *Let T be compact and self-adjoint. Then every eigenvalue λ of T is real. For each such λ, let $V_\lambda \subseteq V$ denote the λ-eigenspace. Then V is the Hilbert direct sum $\hat{\oplus}_\lambda V_\lambda$, i.e., the closure of the algebraic direct sum $\oplus_\lambda V_\lambda$. Moreover, for each $\varepsilon > 0$, the space $\oplus_{\lambda:|\lambda|\geq\varepsilon} V_\lambda$ is finite-dimensional. In particular, if V is nonzero, then T has an eigenvector.*

Proof sketch. Let's note first that the finite-dimensionality of the spaces $\oplus_{\lambda:|\lambda|\geq\varepsilon} V_\lambda$ is automatic: if any such space were infinite-dimensional, then we could find an infinite sequence of unit vectors v_n for which $\|Tv_n\| \geq \varepsilon$, contrary to the assumed compactness of T.

Turning to the main point of the proof, let W denote the orthogonal complement of $\oplus_\lambda V_\lambda$. We know that T acts on W and has no eigenvectors in W, while our task is to show that $W = \{0\}$. The main point is thus to show that any compact self-adjoint operator on a nonzero Hilbert space admits an eigenvector. See for instance §9.2 of my "Lie Groups" notes on the course homepage. $\qquad\square$

Recall also from linear algebra the following consequence of (e.g.) the Jordan normal form:

Lemma 5. *Any linear operator on a nonzero finite-dimensional (complex) vector space has an eigenvector.*

1.6. **Linear algebra.** Let V, V_1, V_2 be finite-dimensional vector spaces. We can then construct some additional vector spaces:

(1) The dual space $V^* = \mathrm{Hom}(V, \mathbb{C})$ consisting of linear functionals $\ell : V \to \mathbb{C}$. For $\ell \in V^*$ and $v \in V$, we sometimes write $\langle \ell, v \rangle = \ell(v)$ for the natural pairing.

(2) The conjugate space $\overline{V}$. By definition, this is a set equipped with a bijection $V \to \overline{V}$, denoted $v \mapsto \overline{v}$. We define the vector space structure on $\overline{V}$ by requiring that $v \mapsto \overline{v}$ commute with addition (i.e., $\overline{v_1} + \overline{v_2} = \overline{v_1 + v_2}$) and intertwine scalar multiplication with conjugation, thus for $\lambda \in \mathbb{C}$,

$$\lambda \overline{v} := \overline{(\overline{\lambda} v)}.$$

(3) The space $\mathrm{Hom}(V_1, V_2)$ of linear maps $T : V_1 \to V_2$.

(4) The space $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ of linear operators $T : V \to V$.

(5) The direct sum $V_1 \oplus V_2$, for which linear maps $f : V_1 \oplus V_2 \to V$ are in natural bijection with pairs (f_1, f_2) of linear maps $f_i : V_i \to V$.

(6) The tensor product $V_1 \otimes V_2$, for which linear maps $f : V_1 \otimes V_2 \to V$ are in natural bijection with bilinear maps $f : V_1 \times V_2 \to V$.

We have an isomorphism

$$V_1^* \otimes V_2 \to \mathrm{Hom}(V_1, V_2)$$

given by

$$\ell_1 \otimes v_2 \mapsto [w_1 \mapsto \ell_1(w_1) v_2]$$

The inverse map is described by taking coefficients of T with respect to bases. In more detail, fix bases $e_1, \ldots, e_m$ of V_1 and $f_1, \ldots, f_n$ of V_2 together with dual bases $e_1^*, \ldots, e_m^*$ of V_1^* and $f_1^*, \ldots, f_n^*$ of V_2^*, so that $\langle e_i^*, e_j \rangle = \langle f_i^*, f_j \rangle = \delta_{ij}$. Then for any $T \in \mathrm{Hom}(V_1, V_2)$ and $v \in V_1$, we have

$$v = \sum_i \langle e_i^*, v \rangle e_i$$

and

$$T e_j = \sum_i \underbrace{\langle f_i^*, T e_j \rangle}_{=:a_{ij}} f_i,$$

so that

$$T v = \sum_{i,j} a_{ij} \langle e_j^*, v \rangle f_i,$$

and thus

$$\mathrm{Hom}(V_1, V_2) \ni T \leftrightarrow \sum_{i,j} a_{ij} e_j^* \otimes f_i \in V_1^* \otimes V_2.$$

In particular, we may identify

$$\mathrm{End}(V) \cong V^* \otimes V.$$

Under this identification,

$$\mathrm{trace} : \mathrm{End}(V) \to \mathbb{C}$$

corresponds to the linear map

$$V^* \otimes V \to \mathbb{C}$$
$$\ell \otimes v \mapsto \langle \ell, v \rangle.$$

2. Introduction

2.1. Representations.

Definition 6. Let G be a topological group and V a topological vector space. A *representation* π of G on V is a group homomorphism

$$\pi : G \to \mathrm{GL}(V)$$

for which the map

$$G \times V \to V$$
$$(g, v) \mapsto \pi(g)v =: gv$$

is continuous.

Rather than saying "(π, V) is a representation of G," we might simply say "let V be a representation of G" or "let G act linearly on V," with the action map $\pi = \pi_V$ defined implicitly and its continuity conditions imposed by default; we will then often denote the action by juxtaposition, $gv := \pi_V(g)v$, as above. Other times we'll say "let π be a representation of G," with the underlying vector space $V = V_\pi$ defined implicitly. Eventually we'll often write simply π both for the action and the underlying vector space, but we'll avoid doing that for now.

Remark 7. Suppose that G is a Lie group and V is finite-dimensional, so that $\mathrm{GL}(V)$ is a Lie group with Lie algebra $\mathrm{End}(V)$. Then a theorem from "Lie groups" implies that π is automatically smooth, hence differentiates to a Lie algebra homomorphism $\mathfrak{g} := \mathrm{Lie}(G) \to \mathrm{End}(V)$.

Example 8.

(1) Let X be a locally compact space equipped with a Radon measure μ. Suppose that G acts on X, preserving μ (i.e., there is a continuous map $G \times X \to X$, satisfying the axioms for a group action, such that for each $g \in G$, the induced map $g : X \to X$ satisfies $g_*\mu = \mu$). Then G acts linearly on $L^2(X, \mu)$, i.e., we have a representation $\pi : G \to \mathrm{GL}(L^2(X, \mu))$. These representations provide some of the most important examples.

In more detail, if we are given a left action denoted "gx," then for $g \in G$ and $f \in L^2(X, \mu)$, we set

$$(\pi(g)f)(x) := f(g^{-1}x), \tag{1}$$

while if we are instead given a right action denoted "xg," then we should take

$$(\pi(g)f)(x) := f(xg). \tag{2}$$

As an exercise, check in either case that $\pi(g_1)\pi(g_2) = \pi(g_1 g_2)$ for all $g_1, g_2 \in G$.

Check for instance for $G = \mathbb{R}$ acting on $X = \mathbb{R}$ by translation, and with μ Lebesgue measure, that the action map

$$\mathbb{R} \times L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

is continuous; on the other hand, for any nonzero $x \in \mathbb{R}$ there exists $f \in L^2(\mathbb{R})$ with $\|f\| = 1$ so that $\|\pi(x)f - f\| \geq 1$ (e.g., take for f a bump function supported in $[-|x|/10, |x|/10]$), so that the map from $\mathbb{R}$ to the space $\mathrm{GL}(L^2(\mathbb{R}))$ equipped with the operator norm is not continuous. These observations motivate the definition of "representation" given above.

Suppose for instance that G is a compact group, equipped with its probability Haar measure dg. Then the action of G on itself by right multiplication preserves dg. The above discussion specialized to $(X, \mu) = (G, dg)$ gives a representation

$$\rho : G \to \mathrm{GL}(L^2(G, dg))$$

given by right translation as in (2). This representation is called the *right regular representation*. We may similarly define the *left regular representation*

$$\lambda : G \to \mathrm{GL}(L^2(G, dg))$$

using left translation as in (1). Note that if G is finite, then $L^2(G, dg)$ is just the space $\mathbb{C}^G$ of functions $f : G \to \mathbb{C}$. Note also that we didn't really require compactness: on any locally compact group, we can define the right (resp. left) regular representation using any right (resp. left) Haar measure.

(2) The *trivial representation* of a group G on the one-dimensional vector space $\mathbb{C}$ is the map $G \to \mathrm{GL}(\mathbb{C}) = \mathrm{GL}_1(\mathbb{C})$ given by $g \mapsto 1$.

(3) The *zero representation* of G on the zero-dimensional vector space $\{0\}$. This representation is unimportant, and will practically never be considered in this course; we mention it for now just to disambiguate it from the (very important) trivial representation.

(4) Most of the classical groups G that one encounters in Lie theory (e.g., $\mathrm{GL}_n(\mathbb{R}), \mathrm{GL}_n(\mathbb{C}), \mathrm{O}(n), \mathrm{U}(n)$ and $\mathrm{Sp}(n)$ for n even) come with a "standard representation" $G \to \mathrm{GL}_n(\mathbb{C}) = \mathrm{GL}(\mathbb{C}^n)$.

(5) For a finite-dimensional vector space V, the group $\mathrm{GL}(V)$ is a Lie group. Its finite-dimensional representations

$$\rho : \mathrm{GL}(V) \to \mathrm{GL}(W)$$

play a special role in the theory: given any representation

$$\pi : G \to \mathrm{GL}(V),$$

we may compose it with ρ to get a new representation

$$\rho \circ \pi : G \to \mathrm{GL}(W).$$

For example, we may take for ρ the determinant representation $\det : \mathrm{GL}(V) \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times$, and form the determinant $\det \circ \pi : G \to \mathbb{C}^\times$ of a representation $\pi : G \to \mathrm{GL}(V)$.

(6) A Lie group G defined over the reals has a Lie algebra $\mathfrak{g}$, which is a real vector space; its complexification $\mathfrak{g}_\mathbb{C} := \mathfrak{g} \otimes_\mathbb{R} \mathbb{C}$ is a complex vector space, and the adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}_\mathbb{R}(\mathfrak{g})$ defines in particular a (complex) representation

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}_\mathbb{C}).$$

(If we start with a Lie group G over the complex numbers, like $\mathrm{GL}_n(\mathbb{C})$, then we get a representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ in the sense of this course without having to complexify.)

(7) The symmetric group $S(n)$ comes with a *standard representation*

$$\pi : S(n) \to \mathrm{GL}_n(\mathbb{C})$$

on $\mathbb{C}^n$, given by permuting the standard basis elements $e_1, \ldots, e_n$: for $g \in S(n)$, we define $\pi(g)e_j := e_{g(j)}$. This is just the representation of

elements of the symmetric group by permutation matrices; for $n = 2$ we get $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, for $n = 3$ we get $\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}$, and so on. We can compose π with the determinant map $\det : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}^\times$ to get the *sign representation*

$$\mathrm{sgn} := \det \circ \pi : S(n) \to \{\pm 1\} \subseteq \mathbb{C}^\times.$$

Equivalently, $\mathrm{sgn}(g) = (-1)^k$ if $g = \tau_1 \cdots \tau_k$, with each τ_j a transposition. We will sometimes use the notation

$$(-1)^g := \mathrm{sgn}(g)$$

for $g \in S(n)$.

(8) Most of the standard operations from linear algebra (e.g., those recalled in §1.6) induce corresponding operations on representations. For instance, given finite-dimensional representations $(\pi, V), (\pi_1, V_1), (\pi_2, V_2)$ of G, we may form:

- the *dual* (or *contragredient*) *representation* (π^*, V^*) on the dual space V^* by setting

$$(\pi^*(g)\ell)(v) := \ell(\pi(g)^{-1}v),$$

so that

$$\langle \pi^*(g)\ell, \pi(g)v \rangle = \langle \ell, v \rangle,$$

- the *conjugate representation* $(\overline{\pi}, \overline{V})$ by

$$\overline{\pi}(g)\overline{v} := \overline{\pi(g)v},$$

- the direct sum $(\pi_1 \oplus \pi_2, V_1 \oplus V_2)$ by

$$g(v_1, v_2) := (gv_1, gv_2),$$

- the tensor product $(\pi_1 \otimes \pi_2, V_1 \otimes V_2)$ by

$$g(v_1 \otimes v_2) := gv_1 \otimes gv_2,$$

and

- the representation $(\mathrm{Hom}(\pi_1, \pi_2), \mathrm{Hom}(V_1, V_2))$ by, for $g \in G$ and $f \in \mathrm{Hom}(V_1, V_2)$,

$$gf := [V_1 \ni v_1 \mapsto \pi_2(g)f(\pi_1(g)^{-1}v_1) \in V_2].$$

2.2. Unitarity.

Definition 9. Let $\pi : G \to \mathrm{GL}(V)$ be a representation as above. Suppose that V is equipped with an inner product $\langle , \rangle$ (e.g., if V is a Hilbert space). We say then that π is *unitary* (with respect to $\langle , \rangle$) if

$$\langle gv_1, gv_2 \rangle = \langle v_1, v_2 \rangle$$

for all $g \in G$ and $v_1, v_2 \in V$; we might equivalently say that the inner product $\langle , \rangle$ is *invariant* for the action of G. We say in general that π is *unitarizable* if there exists an invariant inner product, i.e., an inner product $\langle , \rangle$ on V with respect to which π is unitary.

We can already prove a basic theorem:

Theorem 10. *Let G be a compact group. Let V be a finite-dimensional represen-tation, or a Hilbert space representation. Then V is unitarizable.*

Proof. In the finite-dimensional case, let $\langle,\rangle_0$ be any inner product on V; in the Hilbert case, let it denote the "given" inner product; in either case, note that it defines a continuous map $V \times V \to V$. The idea is to average this inner product using the Haar measure to get an invariant inner product. Turning to details, recall that the action map $G \times V \to V$ is assumed continuous. For v_1, v_2, the function $G \ni g \mapsto \langle gv_1, gv_2\rangle_0$ is then a continuous function on a compact set, hence is bounded, and in particular integrable with respect to the probabilty Haar dg. We set

$$\langle v_1, v_2 \rangle := \int_{g \in G} \langle gv_1, gv_2 \rangle_0 \, dg.$$

It's easy to see that $\langle,\rangle$ defines an inner product; for instance, if $v \neq 0$, then $\langle v, v \rangle_0 > 0$, hence (by continuity) $\langle gv, gv \rangle_0 > 0$ for all g in some neighborhood of the identity element, hence (by regularity of Radon measures) $\langle v, v \rangle > 0$. Using the right-invariance of dg, we verify for $h \in G$ that

$$\langle hv_1, hv_2 \rangle = \int_{g \in G} \langle ghv_1, ghv_2 \rangle_0 \, dg = \int_{g \in G} \langle gv_1, gv_2 \rangle_0 \, dg = \langle v_1, v_2 \rangle.$$

Thus the representation V is unitary with respect to the inner product $\langle,\rangle$. $\qquad\square$

2.3. Morphisms, equivariant maps, intertwiners, isomorphisms, equiva-lences, and so on.

Definition 11. Let (π_1, V_1) and (π_2, V_2) be representations of G. By a *morphism* of representations $\phi : V_1 \to V_2$ we will mean a continuous linear map that is compatible with the given representations in the sense that $\pi_2(g)\phi(v_1) = \phi(\pi_1(g)v_1)$ for all $v_1 \in V_1$ and $g \in G$. The following phrases will be used interchangeably with "morphism":

- G-equivariant map, or simply equivariant map.
- intertwining operator.

We denote by

$$\mathrm{Hom}_G(V_1, V_2)$$

the set of such ϕ. This set is a vector space.

This definition endows the set of representations of a given group G with the structure of a category.

An *isomorphism* (or sometimes "equivalence", etc.) of representations is a mor-phism that admits a two-sided inverse morphism; in the case of finite-dimensional representations, an isomorphism is the same thing as a bijective morphism.

Example 12.

(1) For finite-dimensional representations V_1, V_2 of G, the isomorphism of vec-tor spaces $V_1^* \otimes V_1 \cong \mathrm{Hom}(V_1, V_2)$ (as recalled in §1.6) defines an isomor-phism of representations.

(2) Let $G = \mathbb{Z}/n\mathbb{Z}$; thus G is the finite cyclic group of order n. Take for (π_1, V_1) the right regular representation on $V_1 = \mathbb{C}^G$, thus

$$\pi_1(g)f(x) = f(x + g).$$

Take $V_2 = \mathbb{C}^n$, with π_2 the representation assigning to each $g \in G$ the diagonal matrix $\pi_2(g)$ with entries $1, e(g/n), e(2g/n), e(3g/n), \ldots, e((n -$

$1)g/n)$, where $e(x) := e^{2\pi i x}$. Then (π_1, V_1) and (π_2, V_2) are equivalent representations; an isomorphism is given by the finite Fourier transform

$$\phi : V_1 \to V_2$$

$$\phi(f) := (\hat{f}(0), \hat{f}(1), \dots, \hat{f}(n-1)),$$

$$\hat{f}(x) := \sum_{y \in \mathbb{Z}/n} f(y)e(-xy/n).$$

[2]

We pause to introduce, for any representation (π, V), the notation

$$V^G := \{v \in V : \pi(g)v = v \text{ for all } g \in G\}$$

for the *G-fixed subspace* of V. Recall that we have defined for any representations V_1, V_2 a representation $\operatorname{Hom}(V_1, V_2)$; it thus makes sense to speak of the fixed subspace of the latter representation, and we verify readily from the definitions that

$$\operatorname{Hom}(V_1, V_2)^G = \operatorname{Hom}_G(V_1, V_2).$$

We defined earlier the dual of any finite-dimensional representation. More generally, for any representation for which the underlying vector space is given the structure of a Hilbert space, we define the dual representation using the *continuous* dual, which is then itself a Hilbert space.

Lemma 13. *Let (π, V) be a representation of G such that*
  - *V is a Hilbert space $(V, \langle, \rangle)$, and*
  - *π is unitary.*

Then π is isomorphic to its conjugate dual, and the conjugate and dual of π are isomorphic to each other:

$$\pi \cong \overline{\pi}^*, \quad \overline{\pi} \cong \pi^*.$$

Proof. Since V is a Hilbert space, we have (by what is sometimes called the "Riesz representation theorem") an isomorphism of vector spaces

$$V \to \overline{V}^*$$

$$v \mapsto \langle v, \cdot \rangle.$$

Since π is unitary, this isomorphism is equivariant. This establishes the first isomorphism; the second is obtained similarly. $\square$

2.4. **Reduction and decomposition.**

Definition 14. Let (π, V) be a representation of G. A *closed invariant subspace* $W \subseteq V$ is a closed subspace such that $\pi(g)W \subseteq W$ for all $g \in G$. We note that:
  - A subspace $W \subseteq V$ is closed if and only if the quotient space V/W, equipped with the quotient topology, is Hausdorff. If V is finite-dimensional, then every subspace W is closed. We will practically never discuss non-closed subspaces in this course.
  - The condition $\pi(g)W \subseteq W$, applied both to g and g^{-1}, implies that in fact $\pi(g)W = W$.

---

[2]End of lecture #1, Tuesday, 19 Feb 2019

To each closed invariant subspace $W \subseteq V$ we may associate a *subrepresentation* $\pi : G \to \mathrm{GL}(W)$ and a *quotient representation* $\pi : G \to \mathrm{GL}(V/W)$. We will often use "subrepresentation" as a synonym for "closed invariant subspace."

For instance, if $\dim(W) = 2$ and $\dim(V) = 5$, then we can extend a basis e_1, e_2 for W to a basis $e_1, \ldots, e_5$ for V, and the matrix entries of our representation expressed in terms of this basis look like

$$\pi(g) = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix}.$$

The upper-left 2×2 block corresponds to the subrepresentation on W, the lower-right 3×3 block to the quotient representation on V/W.

Example 15.   (1) For any representation (π, V) of G, the fixed subspace $V^G \subseteq V$ is a closed invariant subspace.
   (2) If G acts in a measure-preserving fashion on some measured space (X, μ) of finite volume, then the representation $V = L^2(X, \mu)$ contains the closed invariant subspace $\mathbb{C}$ consisting of constant functions.

Definition 16. Let (π, V) be a nonzero representation (thus V is not the zero-dimensional space $\{0\}$; it might be the one-dimensional trivial representation).

We say that V is *reducible* if there exists a closed invariant subspace $W \subseteq V$ with $W \neq \{0\}, V$. We say otherwise that V is *irreducible*; this means that $\{0\}$ and V are the only closed invariant subspaces.

Example 17. "Most" V considered previously are reducible. If $\dim(V) = 1$, then V is irreducible.

The irreducible representations are a bit like the prime numbers, with closed invariant subspaces playing the role of divisors and the zero space $\{0\}$ a bit like the unit element 1.

Theorem 18 (Schur's lemma). *Let (π_1, V_1) and (π_2, V_2) be irreducible finite-dimensional representations of some group G. Then*

$$\dim \mathrm{Hom}_G(V_1, V_2) = \begin{cases} 1 & \text{if } V_1 \cong V_2, \\ 0 & \text{otherwise,} \end{cases}$$

where "$V_1 \cong V_2$" means "isomorphic as representations of G." If moreover $V = V_1 = V_2$, then $\mathrm{End}_G(V) := \mathrm{Hom}_G(V, V)$ is the space $\mathbb{C}\,\mathrm{id}$ of scalar multiples of the identity $\mathrm{id} : V \to V$.

Proof. Let ϕ be a nonzero element of $\mathrm{Hom}_G(V_1, V_2)$. Then $\mathrm{image}(\phi) \subseteq V_2$ and $\ker(\phi) \subseteq V_1$ are invariant subspaces. Since $\phi \neq 0$, we must have $\mathrm{image}(\phi) \neq 0$ and $\ker(\phi) \neq V_1$. Thus by the irreducibility of V_1, V_2, we have $\mathrm{image}(\phi) = V_1$ and $\ker(\phi) = 0$. Thus ϕ is an isomorphism. In particular, if V_1, V_2 are non-isomorphic, then $\mathrm{Hom}_G(V_1, V_2) = \{0\}$. If V_1 and V_2 are isomorphic, then the spaces $\mathrm{Hom}_G(V_1, V_2) \cong \mathrm{Hom}_G(V_1, V_1)$ are isomorphic via composition with an isomorphism $V_1 \to V_2$, so in particular $\dim \mathrm{Hom}_G(V_1, V_2) = \dim \mathrm{Hom}_G(V_1, V_1)$; we thereby reduce to establishing the final assertion concerning $\phi \in \mathrm{End}_G(V)$. (**Note:**

this last step was treated incorrectly in lecture!) Since V is a nonzero finite-dimensional vector spaces over the complex numbers, the operator ϕ has an eigenvalue λ. Let V_λ denote the corresponding eigenspace. For any $g \in G$ and $v \in V_\lambda$, we have

$$\phi(\pi(g)v) = \pi(g)\phi(v) = \pi(g)\lambda v = \lambda \pi(g)v, \tag{3}$$

thus $\pi(g)v \in V^\lambda$, i.e., V^λ is a nonzero invariant subspace; since V is irreducible, it follows that $V^\lambda = V$, and so $\phi = \lambda \cdot \mathrm{id}$, as required. $\qquad\square$

Remark 19. Inserted to keep numbering consistent with the numbering that I messed up in lecture; maybe I'll think of something clever to put here later.

Remark 20. Same.

Corollary 21. *Suppose that G is abelian and (π, V) is finite-dimensional and irreducible. Then $\dim(V) = 1$.*

Proof. Let $h \in G$. Then for all $g \in G$, $\pi(h)\pi(g) = \pi(hg) = \pi(gh) = \pi(g)\pi(h)$. Thus $\pi(h) \in \mathrm{End}_G(V) = \mathbb{C}\,\mathrm{id}$, and so $\pi(h)$ stabilizes every line in V. Thus any line in V is a nonzero invariant subspace; by irreducibility, any such line is equal to V, and so $\dim(V) = 1$. $\qquad\square$

As a matter of notation, we now define for a compact group G the set

$\mathrm{Irr}(G) := \{\text{isomorphism classes of finite-dimensional irreducible representations } \pi = (\pi, V_\pi) \text{ of } G\}$.

Theorem 22. *Let G be compact and (π, V) finite-dimensional. Then there exists for each $\sigma = (\sigma, W_\sigma) \in \mathrm{Irr}(G)$ a nonnegative integer $n(\sigma)$ so that we have an isomorphism of representations*

$$V \cong \oplus_{\sigma \in \mathrm{Irr}(G)} W_\sigma^{\oplus n(\sigma)}.$$

(Here $W^{\oplus n} := W \oplus \cdots \oplus W$, with n copies.) The $n(\sigma)$ are determined uniquely by V.

Proof. Since G is compact, we may equip V with an invariant inner product $\langle,\rangle$. We may assume also that V is nonzero, since the conclusion of the theorem holds in that case. Then there exist nonzero invariant subspaces U_1 of V, such as V itself. Choose a *minimal* such subspace U_1. Then the orthogonal complement $U_1^\perp$ is an invariant subspace, since for $v \in U_1^\perp$ and $g \in G$, we have for each $u \in U_1$ that likewise $g^{-1} \in U_1$, and so

$$\langle gv, u \rangle = \langle v, g^{-1}u \rangle = 0.$$

By choosing (if possible) another minimal nonzero invariant subspace $U_2 \le U_1^\perp$ and inducting on dimension (note that $\dim(U_1^\perp) < \dim(V) < \infty$), we see that we may write V as the (orthogonal) direct sum

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_n,$$

where each U_j is an irreducible invariant subspace. We now group the U_j according to their isomorphism class, giving a partition of $\{1, \ldots, n\}$ by $\mathrm{Irr}(G)$; this gives the required decomposition.

The uniqueness of the $n(\sigma)$ may be deduced as in the proof of the Jordan–Hölder theorem; we will give an alternative proof later. $\qquad\square$

This last result suggests two problems for a given G:
- Determine the set $\mathrm{Irr}(G)$ explicitly.

- Explicitly decompose "interesting" V in the above sense.

2.5. **Characters.** Throughout this section we take G compact and (π, V) finite-dimensional.

Definition 23. The *character* of π is the function

$$\chi_\pi : G \to \mathbb{C}$$

$$g \mapsto \mathrm{trace}(\pi(g)).$$

We sometimes write χ_V for χ_π.

A *class function* $f : G \to \mathbb{C}$ is a function that is constant on conjugacy classes, thus $f(g^{-1}xg) = f(x)$ for all $x, g \in G$. Note that characters are class functions (because trace is conjugation-invariant).

The idea is that, given a linear operator on a finite-dimensional vector space, the only way to linearly assign to that operator a scalar without choosing a basis is to take the trace (or a multiple thereof). We're looking to study isomorphism classes of representations, which suggests looking at their characters. Here are some basic properties:

Lemma 24. *Let π, π_1, π_2 be finite-dimensional representations of the compact group G, as above.*

(1) χ_π *depends only upon the isomorphism class of π.*
(2) $\chi_{\overline{\pi}} = \overline{\chi_\pi}$
(3) $\chi_{\pi_1 \oplus \pi_2} = \chi_{\pi_1} + \chi_{\pi_2}$
(4) $\chi_{\pi_1 \otimes \pi_2} = \chi_{\pi_1} \chi_{\pi_2}$
(5) $\chi_{\pi^*} = \chi_{\overline{\pi}} = \overline{\chi_\pi}$
(6) $\chi_{\mathrm{Hom}(\pi_1, \pi_2)} = \overline{\chi_{\pi_1}} \chi_{\pi_2}$

Proof. The only "tricky" part is that since G is compact, π is unitarizable, and so (by Lemma 13) $\overline{\pi} \cong \pi^*$, thus $\chi_{\pi^*} = \chi_{\overline{\pi}}$. We also use that $\mathrm{Hom}(\pi_1, \pi_2) \cong \pi_1^* \otimes \pi_2$. $\square$

The key to unlocking the power of characters in the case of compact groups is the following identity between the dimension of the fixed subspace and the average value of the character:

Lemma 25. *Define a linear map*

$$p : V \to V$$

$$v \mapsto \int_{g \in G} \pi(g)v \, dg.$$

(Here dg denotes as usual the probability Haar, and we are integrating a function $g \mapsto \pi(g)v$ valued in the finite-dimensional vector space V; this can be defined by choosing a basis and integrating each coordinate, for instance.) Then

$$\dim(V^G) = \mathrm{trace}(p) = \int_{g \in G} \chi_\pi(g) \, dg.$$

Proof. We check readily that $p(V) \subseteq V^G$ and that p restricts to the identity on V^G, so that p is a projection onto V^G; its trace is thus the dimension of that subspace, giving the first identity.

The second identity follows from the linearity of the trace map. (In more detail, let's fix a basis $e_1, \ldots, e_n$ for V and dual basis $e_1^*, \ldots, e_n^*$ and write $\pi_{ij}(g) :=$

$\langle e_j^*, \pi(g)e_i \rangle$ for the coefficients of $\pi(g)$ with respect to that basis. Let $p_{ij} := \langle e_j^*, pe_i \rangle$ denote the coefficients of p. By definition, $p_{ij} = \langle e_j^*, \int_{g \in G} \pi(g)e_i\, dg \rangle$. The coefficients of a vector-valued integral are obtained by integrating the coefficients, i.e., $\langle e_j^*, \int_{g \in G} \pi(g)e_i\, dg \rangle = \int_{g \in G} \langle e_j^*, \pi(g)e_i \rangle\, dg$, thus $p_{ij} = \int_{g \in G} \pi_{ij}(g)\, dg$. Summing over $i = j$ gives $\mathrm{trace}(p) = \sum_i p_{ii} = \sum_i \int_{g \in G} \pi_{ii}(g)\, dg = \int_{g \in G} \sum_i \pi_{ii}(g)\, dg = \int_G \chi_\pi$, as required. $\qquad\square$

Theorem 26. *(1) Let $\pi, \pi' \in \mathrm{Irr}(G)$. Then*

$$\langle \chi_\pi, \chi_{\pi'} \rangle_{L^2(G)} = \begin{cases} 1 & \text{if } \pi \cong \pi', \\ 0 & \text{otherwise.} \end{cases} \qquad (4)$$

Thus $\{\chi_\pi : \pi \in \mathrm{Irr}(G)\}$ is an orthonormal subset of

$$L^2(G)^{\mathrm{class}} := \{\text{class functions in } L^2(G)\}.$$

(2) Let (π, V) be any finite-dimensional representation, with decomposition

$$V = \oplus_{\sigma \in \mathrm{Irr}(G)} W_\sigma^{\oplus n(\sigma)}$$

as before. Then

$$n(\sigma) = \langle \chi_\pi, \chi_\sigma \rangle.$$

In particular, χ_π determines π up to isomorphism.

Proof.   (1) By definition, $\langle \chi_\pi, \chi_{\pi'} \rangle = \int_G \chi_\pi \overline{\chi_{\pi'}}$. We have seen that $\chi_\pi \overline{\chi_{\pi'}} = \chi_{\mathrm{Hom}(\pi', \pi)}$, and that the average value of the latter is the dimension of $\mathrm{Hom}(\pi', \pi)^G = \mathrm{Hom}_G(\pi', \pi)$; the required conclusion now follows from Schur's lemma.

(2) We have $\chi_\pi = \sum_\sigma n(\sigma)\chi_\sigma$, so the conclusion follows from (4).

$\qquad\square$

Theorem 27 (Part of the Peter–Weyl theorem)**.** *Recall that G is compact. The orthonormal subset $\{\chi_\pi : \pi \in \mathrm{Irr}(G)\}$ of $L^2(G)^{\mathrm{class}}$ is in fact an orthonormal basis, i.e., has dense span.*

Proof. We postpone this proof until a bit later in the course; it is not so difficult, but we prefer to compute some actual characters first. $\qquad\square$

Corollary 28. *If G is finite, then $\mathrm{Irr}(G)$ is a finite set whose cardinality is the number of conjugacy classes C of G.*

Proof. In that case $L^2(G)^{\mathrm{class}}$ has a basis given by the characteristic functions 1_C of each such C. $\qquad\square$

2.6. **The connected compact abelian case.** Let's illustrate the content of Peter–Weyl by explicating it completely in the case of a connected compact abelian Lie group. In such a group, conjugation is trivial, so "class function" just means "function;" also, we have seen that any irreducible finite-dimensional representation is one-dimensional, hence may be identified with a group homomorphism $\pi : G \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times$, which is thus the "same" as its character, i.e., $\pi(g)$ and $\chi_\pi(g)$ are the same complex scalar for all $g \in G$. Recall that

$$\mathrm{U}(1) := \{z \in \mathbb{C}^\times : |z| = 1\} \cong \mathbb{R}/\mathbb{Z},$$

$$e(\theta) := e^{2\pi i\theta} \leftarrow\!\shortmid \theta.$$

Lemma 29. *Let G be a compact connected abelian Lie group of dimension $n \geq 0$. Then*

$$G \cong \mathrm{U}(1)^n \cong (\mathbb{R}/\mathbb{Z})^n \cong \mathbb{R}^n/\mathbb{Z}^n.$$

One has a bijection

$$\mathrm{Irr}(G) \leftrightarrow \mathbb{Z}^n$$

given by associating to each $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}^n$ the one-dimensional representation

$$\mathrm{U}(1)^n \ni z = (z_1, \ldots, z_n) \mapsto z^\lambda := z_1^{\lambda_1} \cdots z_n^{\lambda_n} \in \mathrm{U}(1) \subseteq \mathrm{GL}_1(\mathbb{C}). \qquad (5)$$

Proof. We verify first that $G \cong \mathrm{U}(1)^n$:

(1) Set $\mathfrak{g} = \mathrm{Lie}(G)$. Since G is abelian, $\exp : \mathfrak{g} \to G$ is a homomorphism.

(2) Since G is connected and $\exp$ is a homomorphism whose image contains a neighborhood of the identity element of G, it follows that $\exp$ is surjective.

(3) The kernel $\Lambda \subseteq \mathfrak{g}$ of $\exp$ is a subgroup. Using that the exponential map is a local diffeomorphism near the origin, we see that Λ is discrete. (Indeed, if there were a sequence $v_n \in \Lambda$ that converged to some element of Λ, then the sequence of differences $v_{n+1} - v_n$ would converge to 0.)

(4) By general theory on quotients of Lie groups (see e.g. Theorem 157 in my Fall 2016 notes), we have $G \cong \mathfrak{g}/\Lambda$. In particular, Λ is cocompact.

(5) It's not hard to see that for every discrete cocompact subgroup Λ of an n-dimensional Euclidean space $\mathfrak{g}$ there exists an isomorphism $\mathfrak{g} \cong \mathbb{R}^n$ under which Λ identifies with $\mathbb{Z}^n$. (Indeed, induct on n. We may assume that $n \geq 1$. Fix an arbitrary Euclidean norm on $\mathfrak{g}$. Since Λ is discrete, we can find a nonzero element $e_1 \in V$ of minimal norm. We might suppose having normalized our norm so that the norm of e_1 is exactly 1. In any event, the minimality of this norm implies that $\mathbb{R}e_1 \cap \Lambda = \mathbb{Z}e_1$. Let W be a subspace of $\mathfrak{g}$ complementary to $\mathbb{R}e_1$ (e.g., the orthogonal complement), and let $p : V \to W$ be the projection with kernel $\mathbb{R}e_1$. We then have a short exact sequence of $\mathbb{Z}$-modules

$$0 \to \mathbb{Z}e_1 \to \Lambda \to p(\Lambda) \to 0.$$

We claim that $p(\Lambda) \subseteq W$ is a discrete subgroup; it follows then inductively that $p(\Lambda)$ is a finite free $\mathbb{Z}$-module of rank at most $\dim(W) = n - 1$, hence that Λ is a free $\mathbb{Z}$-module of rank at most n; the assumed compactness of $\mathfrak{g}/\Lambda$ then forces the rank to equal n, and so $\Lambda \cong \mathbb{Z}^n$ with respect to some coordinates.

To verify the claim, it suffices to bound from below the norms of nonzero elements of $p(\Lambda)$ (this implies that such elements can't accumulate at the origin, hence neither can their differences, giving the required discreteness). So let $v \in \Lambda$ with $p(v) \neq 0$, i.e.,

$$v \notin \mathbb{R}e_1.$$

The difference $v - p(v)$ then lies in $\mathbb{R}e_1$. We can find an integral multiple me_1 of e_1 that "best approximates" this difference in the sense that

$$v - p(v) - me_1 \in [-1/2, 1/2]e_1.$$

We have $v \neq me_1$, so $v - me_1$, being a nonzero element of Λ, has norm bounded from below by 1 (the minimal norm of any such element). But

elements of $[-1/2, 1/2]e_1$ have norm at most $1/2$. The triangle inequality thus forces $p(v)$ to have norm at least $1/2$.)

Having proved that $G \cong \mathrm{U}(1)^n$, we turn to classifying its irreducible finite-dimensional representations π. By Corollary 21, any such π is one-dimensional, and so may be regarded as a homomorphism

$$\pi : G \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times.$$

Compose this with the isomorphism

$$\mathbb{R}^n/\mathbb{Z}^n \ni \theta = (\theta_1, \ldots, \theta_n) \mapsto e(\theta) := (e(\theta_1), \ldots, e(\theta_n)) \in G$$

to get a morphism of Lie groups

$$\tau : \mathbb{R}^n/\mathbb{Z}^n \ni \theta \mapsto \pi(e(\theta)) \in \mathbb{C}^\times.$$

Differentiate this to obtain a morphism of Lie algebras

$$\mathbb{R}^n \ni \theta \mapsto d\tau(\theta) \in \mathrm{Lie}(\mathbb{C}^\times) = \mathbb{C}.$$

Any such morphism is of the form

$$d\tau(\theta) = \lambda_1\theta_1 + \cdots + \lambda_n\theta_n \text{ for some } \lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n.$$

By the compatibility between Lie group morphisms, Lie algebra morphisms and the exponential map, we have

$$\tau(\theta) = e(\lambda_1\theta_1 + \cdots + \lambda_n\theta_n).$$

Since $\tau(\mathbb{Z}^n) = \{1\}$, we have in particular

$$e(\lambda_1) = \cdots = e(\lambda_n) = 1,$$

and so $\lambda_1, \ldots, \lambda_n \in \mathbb{Z}$, as required.

We note finally that all of the one-dimensional representations of $G = \mathrm{U}(1)^n$ that we have defined are inequivalent; for instance, we verify readily that they (and hence, what amounts to the same, their characters) are orthogonal to one another in $L^2(G)$, thanks to repeated application of the basic identity: for $\ell \in \mathbb{Z}$,

$$\int_{\theta \in [0,1]} e(\ell\theta)\, d\theta = \begin{cases} 1 & \text{if } \ell = 0, \\ 0 & \text{otherwise.} \end{cases}$$

$\square$

The content of Peter–Weyl in the setting of the compact connected abelian Lie groups $\mathrm{U}(1)^n$ is thus that, as λ varies over $\mathbb{Z}^n$, the "trigonometric polynomials" $\mathrm{U}(1)^n \ni z \mapsto z^\lambda$ have dense span in $L^2(\mathrm{U}(1)^n)$. This may be a familiar fact from abelian Fourier analysis. It can be established using (e.g.) the Stone–Weierstrass theorem, or a bit of functional analysis. The proof of the general case of the Peter–Weyl theorem, to be given a bit later, will involve similar arguments.[3]

---

[3]End of lecture #2, Thursday, 21 Feb 2019

3. Character theory of compact unitary groups

3.1. Recap. Before beginning, let's pause to attach some vocabulary to results from last time.

Definition 30. A *torus T* is a compact connected abelian Lie group.

We saw last time that any n-dimensional torus T is isomorphic to $\mathrm{U}(1)^n$, and that its irreducible representations are indexed by $\lambda \in \mathbb{Z}^n$ and given by

$$e^\lambda : \mathrm{U}(n) \to \mathrm{U}(1)$$

$$t = (t_1, \ldots, t_n) \mapsto t^\lambda := (t_1^{\lambda_1}, \ldots, t_n^{\lambda_n}).$$

3.2. Conjugacy classes in $\mathrm{U}(n)$.

Definition 31. The compact unitary group $\mathrm{U}(n)$ is defined by

$$\mathrm{U}(n) := \{g \in \mathrm{GL}_n(\mathbb{C}) : \langle gv, gw \rangle = \langle v, w \rangle \text{ for all } v, w \in \mathbb{C}^n\}, \qquad (6)$$

where $\langle v, w \rangle := \sum_{j=1}^n v_j \overline{w_j}$ denotes the standard inner product. We can also describe it as the group of matrices $g = (g_{ij})$ whose rows (or columns) form an orthonormal basis, i.e.,

$$\mathrm{U}(n) = \{g : \sum_k g_{ik}\overline{g_{jk}} = \delta_{ij}\} = \{g : \sum_k g_{ki}\overline{g_{kj}} = \delta_{ij}\}.$$

From the second description we see in particular that each $|g_{ij}| \leq 1$, hence that $\mathrm{U}(n)$ is compact.

Henceforth set

$$G := \mathrm{U}(n).$$

We're interested in studying (finite-dimensional) representations π of G. As we saw last time, we can do this by studying their characters χ_π. Characters are class functions, so we might get started by recalling what the conjugacy classes in G look like.

Let $T \leq G$ denote the subgroup of diagonal elements. Then

$$T \cong \mathrm{U}(1)^n$$

$$\begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} \leftrightarrow (t_1, \ldots, t_n).$$

Let $W \leq G$ denote the subgroup of permutation matrices, i.e., the image of the permutation representation of the symmetric group discussed previously. Then W acts on T by conjugation, permuting coordinates:

$$w \cdot t := wtw^{-1} = (t_{w^{-1}(1)}, t_{w^{-1}(2)}, \ldots, t_{w^{-1}(n)}).$$

Here in writing $w^{-1}(j)$ we regard w as a permutation of $\{1, \ldots, n\}$. We have $|W| = n!$.

Lemma 32. *Every conjugacy class in G intersects T, and two elements of T are conjugate in G precisely when they have the same W-orbit.*

Proof. This is presumably known from a linear algebra course as the "spectral theorem for unitary operators," but perhaps it won't hurt to recall the proof:

Let $g \in G$. We can then find an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ consisting of eigenvectors for g. (Take for v_1 any eigenvector, normalize it to have norm 1, observe that the unitarity of g implies that g stabilizes the orthogonal complement $v_1^\perp$, take for v_2 any norm one eigenvector of g in $v_1^\perp$, and so on. Alternatively, let Γ denote the closure of the group generated by g; then Γ is a compact abelian group to which Theorem 22 applies. The "two proofs" are basically the same, of course.) Let $h \in G$ have rows $v_1, \ldots, v_n$. Let $e_1, \ldots, e_n$ denote the standard orthonormal basis of $\mathbb{C}^n$. Then $he_j = v_j$, so $t := h^{-1}gh$ is a diagonal element of G, i.e., $t \in T$. This shows that every conjugacy class in G intersects T. If two elements of T lie in the same W-orbit, then they are obviously conjugate in G, because $W \leq G$. Conversely, if two elements of T are conjugate in G, then they have the same multiset of eigenvalues, so we can find a permutation $w \in W$ sending one to the other. $\square$

Corollary 33. *Restriction defines a bijection*

$$\{class\ functions\ on\ G\} \leftrightarrow \{W\text{-}invariant\ functions\ on\ T\}.$$

3.3. **Weight space decompositions of representations of** $\mathrm{U}(n)$**.** In particular, let (π, V) be any finite-dimensional representation of G. Then its character $\chi_\pi : G \to \mathbb{C}$ is a class function, hence determined by its restriction $\chi_\pi|_T$ to T, which is W-invariant. The restriction $\chi_\pi|_T$ may be regarded as the character of the restriction $\pi|_T : T \to \mathrm{GL}(V)$ of the representation π. By our general discussion of representations of tori (Lemma 29) and of compact groups (Theorem 22), we may decompose $\pi|_T$ as a direct sum of irreducible representations $e^\lambda : t \mapsto t^\lambda$ of T, each occurring with some multiplicities $m_\pi(\lambda) \in \mathbb{Z}_{\geq 0}$ (denoted "$n(\sigma)$" in the cited theorem). Of course $\dim(\pi) = \sum_\lambda m_\pi(\lambda)$, so only finitely many of the $m_\pi(\lambda)$ are nonzero. Thus for $t \in T$,

$$\chi_\pi(t) = \sum_\lambda m_\pi(\lambda) t^\lambda.$$

Setting

$$V^\lambda := \{v \in V : \pi(t)v = t^\lambda v \text{ for all } t \in T\},$$

we have

$$\dim V^\lambda = m_\pi(\lambda)$$

and

$$V = \oplus_\lambda V^\lambda.$$

If $m_\pi(\lambda) > 0$ (equivalently, $V^\lambda \neq 0$), we say that λ is a *weight* of π and refer to $m_\pi(\lambda)$ as the *multiplicity* of λ in π, to V^λ as the corresponding *weight space*, and to nonzero elements of V^λ as *weight vectors*.

Some examples will presumably clarify matters. In what follows we identify characters with functions on T as above. We also write, e.g., t_j as shorthand for the map $T \ni t \mapsto t_j$.

Example 34. (1) The trivial representation π has character 1, so the only weight is $\lambda = 0$, with multiplicity 1.
(2) The standard representation $\mathbb{C}^n$ has a basis of weight vectors $e_1, \ldots, e_n$ (the standard basis) with corresponding weights $t_1, \ldots, t_n$ and each of multiplicity 1, so the character is $t_1 + \cdots + t_n$.

(3) The dual $(\mathbb{C}^n)^*$ of the standard representation, with basis of weight vectors $e_1^*, \ldots, e_n^*$ (the dual of the standard basis, i.e., $\langle e_i^*, e_j \rangle = \delta_{ij}$) with corresponding weights $t_1^{-1}, \ldots, t_n^{-1}$ each of multiplicity 1, hence character given by $t_1^{-1} + \cdots + t_n^{-1}$.

(4) The complexified adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}_{\mathbb{R}}(\mathfrak{g}) \to \mathrm{GL}(\mathfrak{g}_{\mathbb{C}})$. Here $\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{gl}_n(\mathbb{C})$ because $\mathfrak{gl}_n(\mathbb{C}) \cong \mathfrak{g} \oplus i\mathfrak{g}$ via the map $x \mapsto (\frac{1}{2}(x - \overline{x}^t), \frac{1}{2}(x + \overline{x}^t))$. We can thus identify $\mathfrak{g}_{\mathbb{C}}$ with the space of $n \times n$ complex matrices, which has a standard basis E_{ij}. The action is described in this optic by $\mathrm{Ad}(g)x = gxg^{-1}$. We have $\mathrm{Ad}(t)E_{ij} = (t_i/t_j)E_{ij}$, so the character is given by $\sum_{i,j} t_i/t_j$, which we may rewrite as $n + \sum_{i \neq j} t_i/t_j$. Thus the trivial character of T is a weight with multiplicity n. The nontrivial weights are indexed by $i \neq j$, occur with multiplicity one, and are given by $\varepsilon_i - \varepsilon_j : t \mapsto t_i/t_j$. (These, the nontrivial weights for the adjoint action, are in general called *roots*.)

We note incidentally that the representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}_{\mathbb{C}})$ is isomorphic to $\mathrm{End}(\mathbb{C}^n)$, i.e., to the tensor product $\mathbb{C}^n \otimes (\mathbb{C}^n)^*$ of the standard representation and its dual; this isomorphism is reflected in the character identity $\sum_{i,j} t_i/t_j = (\sum_i t_i)(\sum_j t_j^{-1})$.

(5) The kth symmetric power $\mathrm{Sym}^k(\mathbb{C}^n)$ of the standard representation has a basis of weight vectors $e_{i_1} \cdots e_{i_k}$ indexed by $i_1 \leq \cdots \leq i_k$, with corresponding weights $t_{i_1} \cdots t_{i_k}$, each occurring with multiplicity one; the character is

$$\sum_{i_1 \leq \cdots \leq i_k} t_{i_1} \cdots t_{i_k}. \tag{7}$$

(6) The kth exterior power $\Lambda^k(\mathbb{C}^n)$ of the standard representation has a basis of weight vectors $e_{i_1} \wedge \cdots \wedge e_{i_k}$ indexed by $i_1 < \cdots < i_k$, with corresponding weights $t_{i_1} \cdots t_{i_k}$, each occurring with multiplicity one; the character is

$$\sum_{i_1 < \cdots < i_k} t_{i_1} \cdots t_{i_k},$$

which is often called the *kth elementary symmetric function*.

(7) The kth tensor power $(\mathbb{C}^n)^{\otimes k}$ of the standard representation has basis of weight vectors $e_{i_1} \otimes \cdots \otimes e_{ik}$ indexed by any $i_1, \ldots, i_k$, with corresponding weights $t_{i_1} \cdots t_{i_k}$. The character is

$$\sum_{i_1, \ldots, i_k} t_{i_1} \cdots t_{i_k} = (t_1 + \cdots + t_n)^k,$$

i.e., the kth power of the character of the standard representation, as we could have predicted from Lemma 24. Note that nontrivial multiplicities $m_\pi(\lambda)$ occur in general, and are given by multinomial coefficients.

Note that in all of these examples, we indeed obtained a W-invariant Laurent polynomial $\sum_\lambda m_\pi(\lambda) t^\lambda$.

3.4. The Weyl integral formula for $\mathrm{U}(n)$.

Enough examples for now. Our goal is to use the orthogonality relations for characters of irreducible representations to derive formulas for them. Since we've seen that the characters assume a particularly simple form when we restrict them to T, we might try first to understand how the integral over G of a class function may be expressed in terms of the restriction of that function to T.

Theorem 35 (Weyl integral formula)**.** *For a continuous class function $f : G \to \mathbb{C}$, we have*

$$\int_G f = \frac{1}{|W|} \int_T |\Delta|^2 f,$$

where

$$\Delta(t) := \prod_{i<j}(t_i - t_j).$$

Here both integrals are taken with respect to probability Haar measures. More generally, for any continuous $f : G \to \mathbb{C}$, we have

$$\int_G f = \frac{1}{|W|} \int_{g \in G} \int_{t \in T} |\Delta(t)|^2 f(gtg^{-1}) \, dt \, dg.$$

This should be thought of as a bit like the formula for integrating in polar coordinates in $\mathbb{R}^3$. We give the proof below. For further reading, see

- Weyl's original treatment (see section V.17 of "Group theory and quantum mechanics" in the course references),
- the treatment given in sections I.5 and IV.1 of BTD (the course reference with those author initials)
- section VIII.5 of Knapp's "Lie groups beyond and introduction,"
- section 6.4 of Rossmann's book,
- and others from the reference list.

Proof. We first review some basics concerning integration on manifolds. Let M be a connected N-manifold, and let ω be a nowhere vanishing differential N-form on M. Then ω defines an orientation and a volume measure $C_c(M) \ni f \mapsto \int_M f \, d\omega$: for any function f supported in an oriented coordinate chart $(x_1, \ldots, x_N)$, we set

$$\int_M f \, d\omega := \int_{x_1, \ldots, x_N} f(x_1, \ldots, x_N) \omega(x_1 \wedge \cdots \wedge x_N) \, dx_1 \cdots dx_N,$$

and then extend this definition to general f via a suitable partition of unity. If (M, ω) and (M', ω') are two such pairs and $\phi : M \to M'$ is a differentiable map, that we may express the pullback $\phi^* \omega'$ as $\det(\phi)\omega$ for some function $\det(\phi) : M \to \mathbb{R}$ (the *determinant* of ϕ with respect to ω and ω'). If ϕ is a diffeomorphism, then we have the change of variables formulas: for $f \in C_c(M')$,

$$\int_{M'} f \, d\omega' = \int_M (f \circ \phi) \cdot |\det(\phi)| \, d\omega. \tag{8}$$

Also, determinant is multiplicative with respect to composition in the obvious sense.

We next recall how to integrate on a compact Lie group using differential forms. Let ω be a nonzero element of $\Lambda^N(\mathfrak{g})^*$. Then, via the correspondence between $\mathfrak{g}$ and left-invariant vector fields on G, we may identify ω with a left-invariant N-form $\tilde{\omega}$ on G. For each $g \in G$, the right translation map $R_g : G \to G$ has some determinant $\det(R_g)$ (defined with respect to $\tilde{\omega}$ both on source and target). We check readily that $g \mapsto \det(R_g)$ defines a continuous homomorphism $G \to \mathbb{R}_+^\times$, but since G is compact, any such map is trivial. Thus $\tilde{\omega}$ is both left and right invariant, and so the corresponding volume measure is a Haar measure. We assume ω normalized so that $\int_G d\tilde{\omega} = 1$; then $f \mapsto \int_G f \, d\tilde{\omega}$ is the probability Haar on G. The two-sided invariance means that for any $g_1, g_2 \in G$, the local chart $g_1 e^x g_2 \mapsto x \in \mathfrak{g}$ at $g_1 g_2$ has determinant one at $g_1 g_2$.

We may similarly realize the probability Haar on T as $f \mapsto \int_T f d\tilde{\alpha}$, where $\tilde{\alpha}$ corresponds to some nonzero element $\alpha \in \Lambda^n(\mathfrak{t})^*$.

Equip $\mathfrak{g}$ with any $\mathrm{Ad}(T)$-invariant scalar product (e.g., by averaging), and let $\mathfrak{g}/\mathfrak{t}$ denote the orthogonal complement to $\mathfrak{t}$. Then $\mathfrak{g} = \mathfrak{t} \oplus \mathfrak{g}/\mathfrak{t}$. The complexification $(\mathfrak{g}/\mathfrak{t})_{\mathbb{C}}$ of $\mathfrak{g}/\mathfrak{t}$ is then an $\mathrm{Ad}(T)$-invariant complement to $\mathfrak{t}_{\mathbb{C}}$ in $\mathfrak{g}_{\mathbb{C}}$; by our discussion of the adjoint representation above, we see that

$$(\mathfrak{g}/\mathfrak{t})_{\mathbb{C}} = \oplus_{i \neq j} \mathbb{C} E_{ij},$$

i.e., it is the space of matrices whose diagonal entries vanish.

There is a unique $\beta \in \Lambda^{N-n}(\mathfrak{g}/\mathfrak{t})^*$ so that $\omega = \alpha \wedge \beta$. For $t \in T$, the operator $\mathrm{Ad}(t)^*$ sends β to $\det(\mathrm{Ad}(t)|_{\mathfrak{g}/\mathfrak{t}})\beta$, but (either by explicit calculation, or using the compactness of T) the determinant in question is trivial, so β is $\mathrm{Ad}(T)^*$-invariant. It thus corresponds to a G-invariant $(N-n)$-form $\tilde{\beta}$ on the $(N-n)$-dimensional manifold G/T, whose corresponding volume form is invariant under left translation by G. Fubini's theorem applied in local coordinates implies that for any $f \in C_c(G)$,

$$\int_G f \, d\tilde{\omega} = \int_{g \in G/T} \left(\int_{t \in T} f(gt) \, d\tilde{\alpha}(t) \right) d\tilde{\beta}(g).$$

(Alternatively, note that the RHS defines a Haar probability measure on G, hence equals the LHS by uniqueness of Haar measures.) Applying this relation with $f = 1$ gives in particular that $\int_{G/T} d\tilde{\beta} = 1$, hence that the integral of a function $G/T \to \mathbb{R}$ with respect to $d\tilde{\beta}$ is the same as the integral of its pullback to G with respect to the probability Haar. We can now restate the desired integral formula as

$$\frac{1}{|W|} \int_{g \in G/T} \left(\int_{t \in T} |\Delta(t)|^2 f(gtg^{-1}) \, d\tilde{\alpha}(t) \right) d\tilde{\beta}(g) = \int_G f \, d\tilde{\omega}. \qquad (9)$$

The basic idea of the proof is that the map

$$\phi : G/T \times T \to G$$

$$(gT, t) \mapsto gtg^{-1}$$

is generically $|W|$-to-1 and has determinant $|\Delta|^2(t)$.

Turning to rigorous details, let's compute $\det(\phi)$ with respect to the differential forms $\tilde{\beta} \times \tilde{\alpha}$ on $G/T \times T$ and $\tilde{\omega}$ on G. The scalar $\det(\phi)(gT, t)$ is the same as the determinant at the origin of the composition

$$\mathfrak{g}/\mathfrak{t} \oplus \mathfrak{t} \xdashrightarrow{(x,y) \mapsto (g \exp(x)T, \exp(y)t)} G/T \times T \xrightarrow{\phi} G \xdashrightarrow{g \exp(x'+y) tg^{-1} \mapsto (x',y)} \mathfrak{g}/\mathfrak{t} \oplus \mathfrak{t},$$

since in this composition the outermost two arrows have determinant one at the relevant argument by construction. (A dashed arrow denotes a partial map, defined on a suitable open.) But since

$$(ge^x)(e^y t)(ge^x)^{-1} \simeq g \exp(x - \mathrm{Ad}(t)x + y) tg^{-1},$$

where $\simeq$ means ignoring terms that depend quadratically or higher upon x and y, we see that the above composition has the same determinant at the origin as the linear map

$$(x, y) \mapsto (x - \mathrm{Ad}(t)x, y).$$

We may compute the determinant of the operator $x \mapsto x - \mathrm{Ad}(t)x$ on $\mathfrak{g}/\mathfrak{t}$ after complexifying and using the explicit basis of weight vectors E_{ij}. We obtain in this

way that

$$\det(\phi)(gT, t) = \prod_{i \neq j}(1 - t_i/t_j) = |\Delta|^2(t). \tag{10}$$

We observe next that the function $|\Delta|^2 : T \to \mathbb{R}_{>0}$ extends to a conjugacy-invariant function on G sending $g \in G$ to the product of the squared magnitudes of the differences between its distinct eigenvalues, with $|\Delta|^2(g) = 0$ precisely when g has a repeated eigenvalue. Let G_{sing} denote the *singular subset* consisting of $g \in G$ having a repeated eigenvalue; it is the zero locus of $|\Delta|^2$, hence a compact subset of G. The function $|\Delta|^2$ is real-analytic (e.g., because it is a polynomial in the coefficients of the characteristic polynomial) and not identically zero on G, so its zero locus G_{sing} has measure zero; moreover, the volume of $\{g \in G : |\Delta|^2 < \varepsilon\}$ tends to zero as $\varepsilon \to 0$. By an approximation argument, we can thus reduce to establishing (9) in the special case that f is supported on $\{g \in G : |\Delta|^2 \geq \varepsilon\}$ for some $\varepsilon > 0$. In particular, we may assume that f is supported on a compact subset of the *regular subset* $G_{\text{reg}} := G - G_{\text{sing}}$ consisting of $g \in G$ having distinct eigenvalues.

Set $T_{\text{reg}} := G_{\text{reg}} \cap T$. We observe that it consists of those $t \in T$ having trivial W-stabilizer. (This turns out to be a special feature of $G = \mathrm{U}(n)$; what's essential is that the subset of $t \in T_{\text{reg}}$ with trivial W-stabilizer has full measure.) We observe that for $t \in T_{\text{reg}}$, its G-centralizer is T (indeed, any centralizing element must stabilize its eigenspaces, which are the one-dimensional spaces $\mathbb{C}e_j$). From this and the fact (Lemma 32) that the conjugates of x in T lie in a single W-orbit, we deduce that $w \cdot (gT, t) := (gw^{-1}T, wtw^{-1})$ defines a simply-transitive action of W on the fiber $\{(gT, t) : gtg^{-1} = x\}$ above $x \in G_{\text{reg}}$. In particular, the fibers all have cardinality $|W|$. On the other hand, we see from (10) that the induced map

$$\phi : G/T \times T_{\text{reg}} \to G_{\text{reg}}$$

has everywhere nonvanishing differential. By the inverse function theorem, it follows that ϕ defines $|W|$-fold covering map. By a partition of unity, we may assume that f is supported in a small connected neighborhood $U \subseteq G_{\text{reg}}$ of some $x \in G_{\text{reg}}$ over which $\phi^{-1}(U) \to U$ is the trivial $|W|$-fold cover. We conclude by applying (8) and (10) on each connected component of $\phi^{-1}(U)$. $\qquad\square$

[4]

3.5. Primer on symmetric polynomials.
We have encountered already several elements of the ring of Laurent polynomials in the variable $t_1, \ldots, t_n$. For their further study, it will be convenient to denote that ring by some letter:

$$L := \mathbb{Z}[t_1^{\pm 1}, \ldots, t_n^{\pm}]$$

As a $\mathbb{Z}$-module, L has the basis consisting of the monomials $e^\lambda : t \mapsto t^\lambda = t_1^{\lambda_1} \cdots t_n^{\lambda_n}$ considered previously, thus

$$L = \oplus_{\lambda \in \mathbb{Z}^n} \mathbb{Z}e^\lambda.$$

As above, we refer to the λ as *weights*.

The symmetric group $W \cong S(n)$ acts on L by $w \cdot f(t_1, \ldots, t_n) := f(t_{w(1)}, \ldots, t_{w(n)})$.

Definition 36. We say that $f \in L$ is *symmetric* if $w \cdot f = f$ for all $w \in W$, and *alternating* if $w \cdot f = (-1)^w f$ for all $w \in W$. We denote by L^{sym} and L^{alt} the respective spaces of symmetric and alternating Laurent polynomials.

---

[4]End of lecture #3, Tuesday, 26 Feb 2019

We have seen that for any finite-dimensional representation π of $G = \mathrm{U}(n)$, the restriction $\chi_\pi|_T$ to T of its character χ_π is a symmetric polynomial. To be pedantic, $\chi_\pi|_T$ is the function $\mathrm{U}(1)^n \to \mathbb{C}$ associated to a unique symmetric polynomial; we will identify polynomials with their associated functions freely in what follows. We will also abbreviate the restriction $\chi_\pi|_T$ simply by χ_π, keeping in mind that the character is determined by its restriction in view of Corollary 33. Subject to these identifications, $\chi_\pi \in L^{\mathrm{sym}}$.

On the other hand, the discriminant polynomial $\Delta(t) = \prod_{i<j}(t_i - t_j)$ that appeared in the Weyl character formula is alternating. One can make this more visible by using the Vandermonde determinant formula $\Delta(t) = \det(t_i^{n-j})_{i,j=1..n}$. We can see that $\Delta \in L^{\mathrm{alt}}$ by recalling that the determinant changes sign when we swap a pair of its rows or columns.

The *monomial symmetric polynomials* are defined for $\lambda \in \mathbb{Z}^n$ by

$$M_\lambda := \sum_{\mu \in W \cdot \lambda} e^\mu.$$

They obviously belong to L^{sym}. Note that we sum over the W-orbit of λ without multiplicity. Note also that $M_\lambda(t)$ is symmetric not only with respect to the argument t, but also with respect to the index λ, i.e., $M_{w \cdot \lambda} = M_\lambda$. For example, if $n = 2$, then $M_{(7,6)}(t) = t_1^7 t_2^6 + t_1^6 t_2^7 = M_{(6,7)}(t)$, while $M_{(7,7)}(t) = t_1^7 t_2^7$. We get a basis for L^{sym} from the M_λ's by considering one λ from each W-orbit on $\mathbb{Z}^n$. It will be convenient to choose an explicit set of representatives:

Definition 37. We say that $\lambda \in \mathbb{Z}^n$ is *dominant* if $\lambda_1 \geq \cdots \geq \lambda_n$, and *strictly dominant* if $\lambda_1 > \cdots > \lambda_n$.

Thus the dominant λ give a set of representatives for the W-orbits on $\mathbb{Z}^n$, and so the M_λ for dominant λ give a basis for L^{sym}.

The *elementary symmetric polynomials* are defined for $k = 1, \ldots, n$ by

$$\sigma_k(t) := \sum_{1 \leq i_1 < \cdots < i_k \leq n} t_{i_1} \cdots t_{i_k}.$$

For example, $\sigma_1 = t_1 + \cdots + t_n$, $\sigma_2 = t_1 t_2 + t_1 t_3 + + \cdots + t_{n-1} t_n$, while $\sigma_n = t_1 \cdots t_n$. We might recall from §3.3 that σ_k is the character of the kth exterior power $\Lambda^n(\mathbb{C}^n)$ of the standard representation $\mathbb{C}^n$. The σ_k together with the inverse of σ_n are well-known to generate L^{sym}; for completeness we state this below and sketch the proof.

Lemma 38. *We have*

$$L^{\mathrm{sym}} = \oplus_{\lambda:\,dominant} \mathbb{Z} M_\lambda = \mathbb{Z}[\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \sigma_n, 1/\sigma_n].$$

Proof. The first identity was noted earlier. For the second identity, the containment "$\supseteq$" is clear; we must verify the reverse containment "$\subseteq$", i.e., that M_λ belongs to the RHS for all dominant λ. Replacing M_λ with $M_\lambda/\sigma_n^{\lambda_n}$ has the effect of replacing λ with $(\lambda_1 - \lambda_n, \lambda_2 - \lambda_n, \ldots, \lambda_{n-1} - \lambda_n, 0)$. We may reduce in this way to the case that $\lambda_n = 0$. Having performed this reduction, we now induct on λ with respect to lexicographical ordering (i.e., ordering first by λ_1, then using λ_2 to break the tie if necessary, and so on). For the base case $\lambda = (0, \ldots, 0)$ we have $M_\lambda = 1 \in \mathbb{Z}$. In general, we observe that

$$M_\lambda(t) = t_1^{\lambda_1} \cdots t_n^{\lambda_n} = t_1^{\lambda_1 - \lambda_2}(t_1 t_2)^{\lambda_2 - \lambda_3} \cdots (t_1 \cdots t_{n-1})^{\lambda_{n-1} - \lambda_n}$$
$$= \sigma_1^{\lambda_1 - \lambda_2} \sigma_2^{\lambda_2 - \lambda_3} \cdots \sigma_{n-1}^{\lambda_{n-1} - \lambda_n}(t) + \cdots,$$

where $\cdots$ denotes an integral linear combination of M_μ taken over dominant μ lexicographically lower than λ. By our inductive hypothesis, $M_\mu/\sigma_n^{\mu_n}$ belongs to $\mathbb{Z}[\sigma_1,\ldots,\sigma_{n-1}]$ for each such μ, hence also $M_\lambda \in \mathbb{Z}[\sigma_1,\ldots,\sigma_{n-1}]$, as required. $\square$

The *monomial alternating polynomials* are defined by

$$A_\lambda := \sum_{w\in W}(-1)^w e^{w\cdot\lambda}.$$

They obviously define elements of L^{alt}. The discriminant Δ arises in this way. Indeed, we may expand the determinental formula for $\Delta(t)$ noted earlier as $\sum_{w\in W}(-1)^w \prod_j t_{w(j)}^{n-j}$ or as $\sum_{w\in W}(-1)^w \prod_i t_i^{n-w(i)}$. Introducing the notation

$$\rho := (n-1, n-2, \ldots, 2, 1, 0) \in \mathbb{Z}^n,$$

so that $t^\rho = t_1^{n-1} t_2^{n-2} \cdots t_{n-2}^2 t_{n-1}$, we obtain $\Delta(t) = \sum_{w\in W}(-1)^w t^{w\cdot\rho}$, that is to say,

$$\Delta = A_\rho.$$

As in the case of monomial symmetric polynomials, the A_λ are alternating not just with respect to their argument but also their index: $A_{w\cdot\lambda} = (-1)^w A_\lambda$. In particular, if $\lambda_i = \lambda_j$ for some $i \neq j$, then $A_\lambda = 0$. For instance, if $n = 2$, then $M_{(7,6)}(t) = t_1^7 t_2^6 - t_1^6 t_2^7 = -M_{(6,7)}(t)$, while $M_{(7,7)}(t) = t_1^7 t_2^7 - t_1^7 t_2^7 = 0$. (To established the required vanishing in general, take for w the transposition swapping i and j, so that $w \cdot \lambda = \lambda$; then $A_\lambda = (-1)^w A_{w\cdot\lambda} = -A_\lambda$, i.e., $2A_\lambda = 0$, and so $A_\lambda = 0$. Alternatively, consider cosets in W of the two-element group generated by w, and note that the sum in the definition of A_λ vanishes over each such coset.) If λ has distinct components, then $A_\lambda \neq 0$; indeed, its lexicographically highest term is e^λ.

As λ runs over the W-orbits of $\lambda \in \mathbb{Z}^n$ having distinct components, the A_λ furnish a basis for L^{alt}. The strictly dominant λ give a system of representatives for such orbits. An equivalent system is given by the $\lambda + \rho$ for dominant λ, noting that the map $\lambda \mapsto \lambda+\rho$ induces a bijection between dominant and strictly dominant weights. Thus

$$L^{\mathrm{alt}} = \oplus_{\lambda:\text{strictly dominant}}\mathbb{Z}A_\lambda = \oplus_{\lambda:\text{dominant}}\mathbb{Z}A_{\lambda+\rho}.$$

Lemma 39. Multiplication by Δ induces an isomorphism of $\mathbb{Z}$-modules

$$L^{\mathrm{sym}} \xrightarrow{f\mapsto\Delta f} L^{\mathrm{alt}}.$$

Proof. The indicated map is clearly defined (because $w \cdot (\Delta f) = (w \cdot \Delta)(w \cdot f) = (-1)^w \Delta f$) and injective, so the content here is that any alternating Laurent polynomial h may be divided by Δ inside the ring of Laurent polynomials. After clearing denominators by multiplying h by a sufficiently large power of $t_1 \cdots t_n$, it suffices to show that any alternating polynomial h is divisible by Δ in the ring $\mathbb{Z}[t_1,\ldots,t_n]$ of ordinary polynomials. For this we induct on n. The case $n \geq 1$ is tautological, because then $\Delta = 1$, so suppose $n \geq 2$. We note (since 2 is not a zerodivisor in $\mathbb{Z}$) that the alternating condition on h implies that it vanishes under any substitution $t_i := t_j$ $(i \neq j)$. (Indeed, from some perspectives it would be better to take this property as the *definition* of alternating.) On the other hand, by division with remainder in the variable t_n, we may write $h = (t_n - t_1)q + r$ where $q \in \mathbb{Z}[t_1,\ldots,t_n]$ and $r \in \mathbb{Z}[t_1,\ldots,t_{n-1}]$, but then substituting $t_n := t_1$ shows that $r = 0$, and so

$h = (t_n - t_1)q$. Since h vanishes under the substitution $t_n := t_2$ and our polynomial rings are integral domains, we see that q vanishes under the substitution $t_n := t_2$. We may thus iterate the above division with remainder argument to see that $h = (t_n - t_1)(t_n - t_2) \cdots (t_n - t_{n-1})q'$ for some $q' \in \mathbb{Z}[t_1, \ldots, t_n]$. We may expand q' in powers of t_n as $\sum a_j t_n^j$, with coefficients $a_j \in \mathbb{Z}[t_1, \ldots, t_{n-1}]$. We may likewise expand h in powers of t_n; each coefficient in this expansion is then an alternating polynomial in the variables $t_1, \ldots, t_{n-1}$, so the same holds for the a_j. By our inductive hypothesis, a_j is divisible by $\prod_{i<j \leq n-1}(t_i - t_j)$. It follows as required that f is divisible by Δ. $\qquad\square$

In particular, the division in the following definition makes sense:

Definition 40. The *Schur polynomial* attached to a dominant $\lambda \in \mathbb{Z}^n$ is

$$s_\lambda := \frac{A_{\lambda+\rho}}{A_\rho} \in L^{\mathrm{sym}}.$$

If t is *regular* in the sense that $t_i \neq t_j$ for $i \neq j$, then

$$s_\lambda(t) = \frac{\det(t_i^{\lambda_j + n - j})}{\prod_{i<j}(t_i - t_j)}. \tag{11}$$

Otherwise, $s_\lambda(t)$ must be understood by continuity, or by first simplifying the above rational function to a polynomial. For instance, when $n = 2$, we have for $\lambda_1 \geq \lambda_2$ that

$$s_{(\lambda_1, \lambda_2)}(t) = \frac{\det \begin{pmatrix} t_1^{\lambda_1 + 1} & t_1^{\lambda_2} \\ t_2^{\lambda_1 + 1} & t_2^{\lambda_2} \end{pmatrix}}{t_1 - t_2} = \frac{t_1^{\lambda_1 + 1} t_2^{\lambda_2} - t_1^{\lambda_2} t_2^{\lambda_1 + 1}}{t_1 - t_2}$$
$$= t_1^{\lambda_1} t_2^{\lambda_2} + t_1^{\lambda_1 - 1} t_2^{\lambda_2 + 1} + \cdots + t_1^{\lambda_2} t_2^{\lambda_1}.$$

Since the $A_{\lambda+\rho}$ give a basis for L^{alt}, we deduce:

Lemma 41. The s_λ give a basis for L^{sym}:

$$L^{\mathrm{sym}} = \oplus_{\lambda:dominant} \mathbb{Z} s_\lambda.$$

We'll need later the computation of the values of the Schur polynomials at the identity element $t = 1 = (1, \ldots, 1) \in \mathrm{U}(1)^n$:

Lemma 42. Let $\lambda \in \mathbb{Z}^n$ be dominant. Temporarily abbreviate $\lambda' := \lambda + \rho$, i.e., $\lambda'_j := \lambda_j + n - j$. Then

$$s_\lambda(1) = \prod_{i<j} \frac{\lambda'_i - \lambda'_j}{j - i}.$$

In particular, $s_\lambda(1) > 0$.

Proof. A direct attempt via (11) gives "0/0," so we need to take a limit along some sequence of regular elements t tending to 1. A clever choice is given by $t_j := z^{n-j}$, where $z \in \mathrm{U}(1)$ traverses a sequence tending to 1. The numerator then simplifies to another Vandermonde determinant:

$$\det(t_i^{\lambda_j + n - j}) = \det(z^{(n-i)\lambda'_j}) = \prod_{i<j}(z^{\lambda'_i} - z^{\lambda'_j}).$$

Using that $z^a - z^b \sim \log(z)(a - b)$, we arrive at the required formula. $\qquad\square$

We record some inner product formulas:

Lemma 43. *Let* $\langle,\rangle$ *denote the inner product in* $L^2(T)$ *taken with respect to the probability Haar, so that* $\langle e^\lambda, e^\mu \rangle = \delta_{\lambda\mu}$. *For strictly dominant* λ, μ, *we have*

$$\langle \Delta s_\lambda, \Delta s_\mu \rangle = \langle A_{\lambda+\rho}, A_{\mu+\rho} \rangle = |W|\delta_{\lambda\mu}.$$

Proof. Immediate. $\square$

3.6. Weyl character and dimension formulas for $\mathrm{U}(n)$.

Theorem 44. *Let* (π, V) *belong to the set* $\mathrm{Irr}(\mathrm{U}(n))$ *of (isomorphism classes of) irreducible finite-dimensional representations of* $\mathrm{U}(n)$.

(1) *There exists a unique dominant weight* $\lambda \in \mathbb{Z}^n$ *so that (the restriction to the torus of) the character* χ_π *is equal to the Schur polynomial* s_λ. *We write in this case* $(\pi, V) = (\pi_\lambda, V_\lambda)$. *The weight space* V^λ *is one-dimensional, while* $V^\mu = \{0\}$ *for any lexicographically larger weight* μ.

(2) *Every dominant weight* λ *arises in this way, giving a bijection*

$$\mathrm{Irr}(\mathrm{U}(n)) \leftrightarrow \{dominant\ \lambda \in \mathbb{Z}^n\}.$$

Proof. We first prove part (i). Recall the notation $G := \mathrm{U}(n) \geq T \cong \mathrm{U}(1)^n$. Write $\chi_\pi = \sum_\mu m_\pi(\mu)e^\mu$ as before. Let λ be the lexicographically highest weight for which $m_\pi(\lambda) \neq 0$. Since χ_π is W-invariant, we know that λ is dominant, as otherwise it would be lexicographically smaller than something in its W-orbit. In any event,

$$\chi_\pi = m_\pi(\lambda)e^\lambda + \cdots, \tag{12}$$

where $\cdots$ denotes the contribution from lexicographically smaller weights. Since also $\Delta = e^\rho + \cdots$, it follows that

$$\Delta\chi_\pi = m_\pi(\lambda)e^{\lambda+\rho} + \cdots. \tag{13}$$

Since the Laurent polynomial $\Delta\chi_\pi$ is alternating, we may group its terms into monomial alternating functions, giving

$$\Delta\chi_\pi = m_\pi(\lambda)A_{\lambda+\rho} + \cdots, \tag{14}$$

where $\cdots$ denotes a linear combination of $A_{\mu+\rho}$ taken over dominant μ lexicographically lower than λ. We combine this expansion with the orthogonality relations for characters (Theorem 26), the Weyl integral formula (Theorem 35), and the inner product formulas for monomial alternating functions (Lemma 43) to see that

$$1 = \langle \chi_\pi, \chi_\pi \rangle_G = \frac{1}{|W|}\langle \Delta\chi_\pi, \Delta\chi_\pi \rangle_T \geq \frac{1}{|W|}\langle m_\pi(\lambda)A_{\lambda+\rho}, m_\pi(\lambda)A_{\lambda+\rho} \rangle_T = |m_\pi(\lambda)|^2.$$

But since $m_\pi(\lambda)$ is a positive integer, it follows that $m_\pi(\lambda) = 1$ and also that equality holds in each step. In particular, the remainder terms "$\cdots$" in (14) must vanish identically, i.e., $\Delta\chi_\pi = A_{\lambda+\rho}$, giving the required formula $\chi_\pi = s_\lambda$. The uniqueness of λ is clear from the orthogonality relations for the Schur polynomials.

We give two proofs that every dominant weight λ arises in this way:

(1) Assume for the sake of contradiction that λ does not arise, i.e., that $\chi_\pi \neq s_\lambda$ for all $\pi \in \mathrm{Irr}(G)$. We may uniquely extend the W-invariant function s_λ on T to a class function on G. For each $\pi \in \mathrm{Irr}(G)$, we may write $\chi_\pi = s_{\lambda'}$ for some dominant $\lambda' \neq \lambda$. By the orthogonality relations for the Schur

polynomials, it follows then that s_λ is orthogonal to χ_π in $L^2(G)$. But this contradicts the Peter–Weyl theorem (Theorem 27).

(2) Let W denote the following representation of G, whose definition may be motivated by the proof of Lemma 38:

$$W := \Lambda(\mathbb{C}^n)^{\otimes \lambda_1 - \lambda_2} \otimes \Lambda^2(\mathbb{C}^n)^{\otimes \lambda_2 - \lambda_3} \otimes \cdots \otimes \Lambda^{n-1}(\mathbb{C}^n)^{\otimes \lambda_{n-1} - \lambda_n} \otimes \Lambda^n(\mathbb{C}^n)^{\lambda_n}.$$

(Note that $\Lambda^n(\mathbb{C}^n)$ is the one-dimensional determinant representation, so it makes sense to raise it to any integral power λ_n, giving the one-dimensional representation $\mathrm{U}(n) \ni g \mapsto \det(g)^{\lambda_n} \in \mathrm{U}(1)$ denoted above by $\Lambda^n(\mathbb{C}^n)^{\lambda_n}$.) Thus, as before,

$$\chi_W = \sigma_1^{\lambda_1 - \lambda_2} \sigma_2^{\lambda_2 - \lambda_3} \cdots \sigma_{n-1}^{\lambda_{n-1} - \lambda_n} \sigma_n^{\lambda_n} = e^\lambda + \cdots,$$

where $\cdots$ denotes the contribution of lexicographically smaller weights. On the other hand, we may decompose W into irreducible subspaces (Theorem 26), taking into account the recently-established description of $\mathrm{Irr}(\mathrm{U}(n))$:

$$W = \oplus_{\text{dominant } \mu} V_\mu^{\oplus n(\mu)}.$$

(By convention, here we sum only over those μ which we as yet know to arise from $\mathrm{Irr}(\mathrm{U}(n))$.) At the level of characters, this decomposition reads $\chi_W = \sum_\mu n(\mu)\chi_\mu$. Multiplying through by $\Delta = e^\rho + \cdots$ and taking into account the formula $\Delta \chi_\mu = A_{\mu+\rho} = e^{\mu+\rho} + \cdots$, we obtain

$$e^{\lambda+\rho} + \cdots = n(\lambda)e^{\lambda+\rho} + \cdots.$$

Thus $n(\lambda) = 1$, that is to say, V_λ occurs (exactly once) as a subrepresentation of W. In particular, the representation V_λ exists in the first place, as required.

(A third proof, closely related to the second: if s_λ is orthogonal to the character of every irreducible representation, then it is orthogonal in particular to any of the ring $\mathbb{Z}[\sigma_1, \ldots, \sigma_{n-1}, \sigma_n, 1/\sigma_n]$ generated by the characters of the exterior powers of the standard representation and the inverse of the determinant, but we have seen that this ring is equal to L^{sym}.)

$\square$

We note incidentally that the second proof given above of the surjectivity of the map $\mathrm{Irr}(\mathrm{U}(n)) \to \{\text{dominant } \lambda \in \mathbb{Z}^n\}$ gives an independent proof of the Peter–Weyl theorem in this case. We note also that, with notation as in that proof, we may realize V_λ inside W explicitly, as follows. The proof shows that the weight space $W[\lambda]$ is one-dimensional, hence coincides with the one-dimensional weight space $V_\lambda[\lambda]$. Consequently, V_λ is the representation generated by (i.e., spanned by the G-orbit of) any nonzero element of $W[\lambda]$, such as $e_1^{\otimes \lambda_1 - \lambda_2} \otimes (e_1 \wedge e_2)^{\otimes \lambda_2 - \lambda_3} \otimes \cdots$.

This result gives a satisfying description of the character theory of $\mathrm{U}(n)$. It has some immediate applications:

(1) A formula for the multiplicity in $m_{\pi_\lambda}(\mu)$ of the weight μ in the representation π_λ, as the coefficient of e^μ in $s_\lambda = A_{\lambda+\rho}/\Delta$. One way to "evaluate" this further is to write $\Delta(t) = t^\rho \prod_{i<j}(1 - t_j/t_i)$ and expand $(1 - t_j/t_i)^{-1}$ as a geometric series (look up the *Kostant partition formula*).

(2) A formula for the coefficients of the decomposition of a tensor product of irreducible representations of $\mathrm{U}(n)$: writing

$$V_{\lambda'} \otimes V_{\lambda'} \cong \oplus_\lambda n(\lambda) V_\lambda,$$

we have $\chi_{\lambda'}\chi_{\lambda''} = \sum_{\lambda} n(\lambda)\chi_{\lambda}$, hence

$$\sum_{\lambda} n(\lambda)A_{\lambda+\rho} = \Delta\chi_{\lambda'}\chi_{\lambda''} = \frac{A_{\lambda'+\rho}A_{\lambda''+\rho}}{\Delta}.$$

We may then determine $n(\lambda)$ as the coefficient of $e^{\lambda+\rho}$ on both sides. (Look up the *Steinberg multiplicity formula*.)

(3) A proof of the irreducibility of several of the representations considered in §3.3, such as $\mathrm{Sym}^k(\mathbb{C}^n)$ or $\Lambda^k(\mathbb{C}^n)$. (A direct "algebraic" proof of their irreducibility is not very difficult, so the present argument should be regarded as an alternative "analytic" proof.) For example, consider $V = \mathrm{Sym}^k(\mathbb{C}^n)$. By "stars and bars," we have (TODO: fix binomial coefficients) On the other hand, $\lambda = (k, 0, \ldots, 0)$ is the lexicographically highest weight of V. Thus λ is lexicographically highest among the dominant weights μ contributing to the decomposition $\chi_V = \sum_{\mu} n(\mu)\chi_{\mu}$, and so V_{λ} must occur as an irreducible subrepresentation of V. In particular, $\dim(V_{\lambda}) \leq \dim(V)$. On the other hand, a short calculation with the dimension formula gives readily the "numerical coincidence" $\dim(V_{\lambda}) = \dim(V)$ (see the homework), from which we deduce that $V \cong V_{\lambda}$; in particular, V is irreducible. A similar argument applies to the exterior powers.

5

3.7. Some groups closely related to $\mathrm{U}(n)$. Set

$$\mathrm{SU}(n) := \{g \in \mathrm{U}(n) : \det(g) = 1\}.$$

Let

$$Z := \{\begin{pmatrix} z & & \\ & \cdots & \\ & & z \end{pmatrix} \in \mathrm{U}(n) : z \in \mathrm{U}(1)\} \cong \mathrm{U}(1) \qquad (15)$$

denote the center of $\mathrm{U}(n)$, and let

$$(n) := \mathrm{U}(n)/Z$$

denote the corresponding quotient. Observe that

- given a representation of (n), we obtain a representation of $\mathrm{U}(n)$ by pull-back, i.e., by composing with the projection $\mathrm{U}(n) \twoheadrightarrow (n)$, and
- given a representation of $\mathrm{U}(n)$, we obtain a representation of $\mathrm{SU}(n)$ by restriction, i.e., by composing with the inclusion $\mathrm{SU}(n) \to \mathrm{U}(n)$.

Theorem 45. *The operations of pullback and restriction just described preserve irreducibility, inducing an injective map*

$$\mathrm{Irr}((n)) \hookrightarrow \mathrm{Irr}(\mathrm{U}(n))$$

and a surjective map

$$\mathrm{Irr}(\mathrm{U}(n)) \twoheadrightarrow \mathrm{Irr}(\mathrm{SU}(n)).$$

The latter maps and the bijection $\mathrm{Irr}(\mathrm{U}(n)) \leftrightarrow \{\text{dominant elements of } \mathbb{Z}^n\}$ as in the Weyl character formula are compatible with bijections

$$\mathrm{Irr}(\mathrm{SU}(n)) \leftrightarrow \{\text{dominant elements of } \mathbb{Z}^n/\mathbb{Z}e_0\}, \quad e_0 := (1, \ldots, 1)$$

---

[5]End of lecture #4, Thursday, 28 Feb 2019

and

$$\mathrm{Irr}((n)) \leftrightarrow \{ \textit{dominant elements of } \mathbb{Z}_0^n := \{\lambda \in \mathbb{Z}^n : \sum_j \lambda_j = 0\}\}.$$

Proof. We leave most of this to the homework; here we just record a few remarks that hopefully make the description of what happens under restriction to $\mathrm{SU}(n)$ seem plausible. Observe that if λ, μ are dominant elements of $\mathbb{Z}^n$ differing by a multiple of e_0, say $\lambda = \mu + \ell e_0$, then the corresponding irreducible representations π_λ, π_μ of $\mathrm{U}(n)$ satisfy

$$\pi_\lambda \cong \pi_\mu \otimes \det{}^\ell.$$

Indeed, we may check this by comparing the characters of these representations, and we have

$$\chi_\lambda(t) = \frac{\det(t_i^{\lambda_j+n-j})}{\Delta(t)} = \underbrace{(t_1 \cdots t_n)^\ell}_{\det(t)^\ell} \frac{\det(t_i^{\mu_j+n-j})}{\Delta(t)}.$$

It follows that π_λ and π_μ have isomorphic restrictions to $\mathrm{SU}(n-1)$. $\square$

3.8. The case of $\mathrm{SU}(2)$.

Theorem 46. *We have bijections*

$$\mathrm{Irr}(\mathrm{SU}(2)) \leftrightarrow \mathbb{Z}_{\geq 0} \leftrightarrow \{ \textit{dominant } \lambda = (\lambda_1, \lambda_2) \in \mathbb{Z}^2/\mathbb{Z}(1,1)\}$$

$$\pi_\lambda \leftrightarrow \lambda \leftrightarrow (\textit{the class of } (\lambda, 0))$$

such that the character χ_λ of π_λ is given by

$$\chi_\lambda \begin{pmatrix} t & \\ & t^{-1} \end{pmatrix} = \frac{\det \begin{pmatrix} t^{\lambda+1} & 1 \\ t^{-\lambda-1} & 1 \end{pmatrix}}{t - t^{-1}} = t^\lambda + t^{\lambda-2} + t^{\lambda-4} + \cdots + t^{-\lambda}. \qquad (16)$$

We have $\dim(\pi_\lambda) = \lambda + 1$. We have

$$\pi_\lambda \cong \mathrm{Sym}^\lambda(\mathbb{C}^2) \cong \mathrm{Sym}^\lambda((\mathbb{C}^2)^*), \qquad (17)$$

where $\mathbb{C}^2$ denotes as usual the standard representation of $\mathrm{SU}(2)$. If we let $x, y \in (\mathbb{C}^2)^*$ denote the standard basis elements of the dual given in coordinates $z = (z_1, z_2)$ by $x : z \mapsto z_1, y : z \mapsto z_2$, then $\pi_\lambda \cong \mathrm{Sym}^\lambda((\mathbb{C}^2)^*)$ identifies with the space $\mathbb{C}[x,y]^{(\lambda)}$ of homogeneous polynomials f of degree λ in the variables x and y, with the action of $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SU}(2)$ given by $g \cdot f(x,y) = f((x,y)g) = f(xa + yc, xb + yd)$; a basis of weight vectors for $T = \left\{ \begin{pmatrix} t & \\ & t^{-1} \end{pmatrix} \right\} \leq \mathrm{SU}(2)$ is given by

$$x^\lambda, x^{\lambda-1}y, x^{\lambda-2}y^2, \ldots, y^\lambda, \qquad (18)$$

with corresopnding weights

$$\begin{pmatrix} t & \\ & t^{-1} \end{pmatrix} \mapsto t^\lambda, t^{\lambda-2}, t^{\lambda-4}, \ldots, t^{-\lambda}.$$

Finally, we have the Clebsch–Gordan rule: for $\lambda, \mu \in \mathbb{Z}_{\geq 0}$,

$$\pi_\lambda \otimes \pi_\mu \cong \pi_{\lambda+\mu} \oplus \pi_{\lambda+\mu-2} \oplus \pi_{\lambda+\mu-4} \oplus \cdots \oplus \pi_{|\lambda-\mu|}. \qquad (19)$$

Proof. Almost every assertion follows immediately from Theorem 45 and the description (7) of the weights of symmetric powers of the standard representation, together with the analogous description of those of its dual. To verify the existence of the isomorphism (19), we just need to check that $\chi_\lambda \chi_\mu = \chi_{\lambda+\mu} + \cdots + \chi_{|\lambda-\mu|}$, which is easy to derive from (16). $\square$

3.9. **Branching problems.** Given compact groups $H \leq G$ and a representation (π, V) of G, we may form the restricted representation $\pi|_H : H \to \mathrm{GL}(V|_H)$ of H; here $V|_H$ is just V, but regarded as a representation of H. The branching problem is to describe this restriction in terms of π.

We assume henceforth that π is irreducible. Its restriction is then "typically" not irreducible, but may in any event be decomposed as

$$V|_H \cong \oplus_{\sigma \in \mathrm{Irr}(H)} W_\sigma^{\oplus n_\pi(\sigma)}$$

for some nonnegative integers $n_\pi(\sigma)$, called "branching coefficients." The problem is to describe these.

We start with some examples that are either basic or have already been discussed implicitly.

(1) If $H = G$, then $n_\pi(\sigma)$ is 1 if $\sigma \cong \pi$ and 0 otherwise.
(2) If $H = \{1\}$ is the trivial subgroup, then $\mathrm{Irr}(H)$ is a singleton consisting of the trivial representation $\sigma_{\mathrm{trivial}}$, and we have $n_\pi(\sigma_{\mathrm{trivial}}) = \dim(\pi)$.
(3) If $G = \mathrm{U}(n)$ and $H \cong \mathrm{U}(1)^n$ is the diagonal subgroup, then σ is the one-dimensional representation $\sigma = \sigma_\mu := e^\mu : t \mapsto t^\mu = t_1^{\mu_1} \cdots t_n^{\mu_n}$ attached to some weight $\mu \in \mathbb{Z}^n$. The branching coefficient $n_\pi(\sigma)$ is just the weight multiplicity $m_\pi(\mu)$ considered in §3.3 and onwards. These integers are typically > 1 when $n \geq 3$ (cf. homework).
(4) If $G = \mathrm{U}(n)$ and $H = \mathrm{SU}(n)$, then Theorem 45 implies that for each π there is exactly one σ (namely, the restriction of π) for which $n_\pi(\sigma)$ is nonzero, and for that σ, we have $n_\pi(\sigma) = 1$.

The overall theme here is that the larger the subgroup H, the more likely the representations of G are to have irreducible restriction to H, or at least to have "mild" branching coefficients. This theme is supported further by the example

$$H := \mathrm{U}(n-1) \hookrightarrow G := \mathrm{U}(n)$$

$$h \mapsto \begin{pmatrix} h & 0 \\ 0 & 1 \end{pmatrix},$$

which we now address. As a matter of notation, for dominant elements $\lambda \in \mathbb{Z}^n$ and $\mu \in \mathbb{Z}^{n-1}$, let us write $\pi_\lambda \in \mathrm{Irr}(\mathrm{U}(n))$ and $\sigma_\mu \in \mathrm{Irr}(\mathrm{U}(n-1))$ for the corresponding irreducibles, and $n_\lambda(\mu) := n_{\pi_\lambda}(\sigma_\mu)$ for the branching coefficient. We say that μ *interlaces* λ, denoted $\mu \prec \lambda$, if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n.$$

Theorem 47. *We have $n_\lambda(\mu) = 0$ unless $\mu \prec \lambda$, in which case $n_\lambda(\mu) = 1$.*

Proof. We must verify that

$$\pi_\lambda|_{\mathrm{U}(n-1)} \cong \oplus_{\mu \prec \lambda} \sigma_\mu. \tag{20}$$

It suffices to check that the representations of $\mathrm{U}(n-1)$ appearing on both sides have the same character. It suffices to compare their characters on the representatives $t = \mathrm{diag}(t_1, \ldots, t_{n-1}, 1)$ for the conjugacy classes in $\mathrm{U}(n-1) \leq \mathrm{U}(n)$. We'll undertake

this comparison in detail when $n = 3$; the general case is similar, but with many occurrences of $(\cdots)$. By the character formula, our task is to show that

$$\frac{1}{(t_1 - t_2)(t_1 - 1)(t_2 - 1)} \det \begin{pmatrix} t_1^{\lambda_1+2} & t_1^{\lambda_2+1} & t_1^{\lambda_3} \\ t_2^{\lambda_1+2} & t_2^{\lambda_2+1} & t_2^{\lambda_3} \\ 1 & 1 & 1 \end{pmatrix} = \frac{1}{t_1 - t_2} \sum_{\mu \prec \lambda} \det \begin{pmatrix} t_1^{\mu_1+1} & t_1^{\mu_2} \\ t_2^{\mu_1+1} & t_2^{\mu_2} \end{pmatrix}.$$

We simplify the 3×3 determinant using elementary column operations, replacing its columns a, b, c with $a - b, b - c, c$. That determinant then simplifies to

$$\det \begin{pmatrix} t_1^{\lambda_1+2} - t_1^{\lambda_2+1} & t_1^{\lambda_2+1} - t_1^{\lambda_3} & t_1^{\lambda_3} \\ t_2^{\lambda_1+2} - t_2^{\lambda_2+1} & t_2^{\lambda_2+1} - t_2^{\lambda_3} & t_2^{\lambda_3} \\ 0 & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} t_1^{\lambda_1+2} - t_1^{\lambda_2+1} & t_1^{\lambda_2+1} - t_1^{\lambda_3} \\ t_2^{\lambda_1+2} - t_2^{\lambda_2+1} & t_2^{\lambda_2+1} - t_2^{\lambda_3} \end{pmatrix}.$$

Using that the determinant is linear in the rows, and expanding out the definition of "$\mu \prec \lambda$," our task reduces to checking that

$$\det \begin{pmatrix} \frac{t_1^{\lambda_1+2} - t_1^{\lambda_2+1}}{t_1 - 1} & \frac{t_1^{\lambda_2+1} - t_1^{\lambda_3}}{t_1 - 1} \\ \frac{t_2^{\lambda_1+2} - t_2^{\lambda_2+1}}{t_2 - 1} & \frac{t_2^{\lambda_2+1} - t_2^{\lambda_3}}{t_2 - 1} \end{pmatrix} = \sum_{\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \lambda_3} \det \begin{pmatrix} t_1^{\mu_1+1} & t_1^{\mu_2} \\ t_2^{\mu_1+1} & t_2^{\mu_2} \end{pmatrix}.$$

This identity follows from the linearity of the determinant with respect to columns and geometric series identities such as $\frac{t_1^{\lambda_1+2} - t_1^{\lambda_2+1}}{t_1 - 1} = \sum_{\lambda_1 \geq \mu_1 \geq \lambda_2} t^{\mu_1+1}$. $\qquad\square$

Groups (G, H) with the property that every branching coefficient $n_\pi(\sigma)$ belongs to $\{0, 1\}$ are called *strong Gelfand pairs*; we may discuss them in more detail later.

3.10. **Gelfand–Tsetlin bases.** We now describe briefly how to deduce from the formulas for the branching coefficients for $U(n-1) \leq U(n)$ a canonical decomposition of any irreducible representation of $U(n)$. Let $\mu \in \mathbb{Z}^n$ be dominant, and $(\pi_\mu, V_\mu) \in \mathrm{Irr}(U(n))$ the corresponding representation. The decomposition (20) implies that for each dominant $\nu \in \mathbb{Z}^{n-1}$ that interlaces μ, there is a unique $U(n-1)$-subrepresentation $V_{\binom{\mu}{\nu}}$ of V_μ of parameter ν, and moreover

$$V_\mu = \oplus_{\nu \prec \mu} V_{\binom{\mu}{\nu}}. \tag{21}$$

The idea is now to iterate this decomposition using the chain of subgroups $U(n-2), U(n-3), \ldots$, all the way down to $U(1)$, whose irreducible representations are all one-dimensional. In preparation for this iteration, we set $\lambda^{(n)} := \mu$ and denote by $\lambda^{(j)}$ a dominant element of $\mathbb{Z}^j$. With this notation, (21) now reads

$$V_\mu|_{U(n-1)} = \oplus_{\lambda^{(n-1)} \prec \lambda^{(n)} = \mu} V_{\binom{\lambda^{(n)}}{\lambda^{(n-1)}}}.$$

We now apply the same reasoning to the restriction to $U(n-2)$ of each representation $V_{\binom{\lambda^{(n)}}{\lambda^{(n-1)}}}$ of $U(n-1)$ arising above, giving

$$V_\mu|_{U(n-2)} = \oplus_{\lambda^{(n-2)} \prec \lambda^{(n-1)} \prec \lambda^{(n)} = \mu} V_{\begin{pmatrix} \lambda^{(n)} \\ \lambda^{(n-1)} \\ \lambda^{(n-2)} \end{pmatrix}}.$$

Proceeding, we eventually obtain a canonical decomposition of V_μ into one-dimensional subspaces:

$$V_\mu = \oplus_{\lambda^{(1)} \prec \lambda^{(2)} \prec \cdots \prec \lambda^{(n)} = \mu} V_{\begin{pmatrix} \lambda^{(n)} \\ \cdots \\ \lambda^{(1)} \end{pmatrix}}.$$

Let's formalize things a bit:

Definition 48. A *Gelfand–Tsetlin pattern of order* $n \geq 1$ is a column vector $\lambda = \begin{pmatrix} \lambda^{(n)} \\ \cdots \\ \lambda^{(1)} \end{pmatrix}$, with $\lambda^{(j)} \in \mathbb{Z}^j$, so that $\lambda^{(1)} \prec \cdots \prec \lambda^{(n)}$. For instance, if $n = 3$, then we can visualize $\mathrm{GT}(n)$ as the space of all triangular arrays

$$\lambda = \begin{pmatrix} \lambda_1^{(3)} & & \lambda_2^{(3)} & & \lambda_3^{(3)} \\ & \lambda_1^{(2)} & & \lambda_2^{(2)} & \\ & & \lambda_1^{(1)} & & \end{pmatrix}$$

satisfying the interlacing conditions

$$\lambda_1^{(3)} \leq \lambda_1^{(2)} \leq \lambda_2^{(3)} \leq \lambda_2^{(2)} \leq \lambda_3^{(3)} \tag{22}$$

and

$$\lambda_1^{(2)} \leq \lambda_1^{(1)} \leq \lambda_2^{(2)}. \tag{23}$$

We denote by $\mathrm{GT}(n)$ the set of all Gelfand–Tsetlin patterns λ of order n, and by

$$\mathrm{GT}(n)_\mu := \{\lambda \in \mathrm{GT}(n) : \lambda^{(n)} = \mu\}$$

the subset of Gelfand–Tsetlin patterns having "top row" equal to μ.

Theorem 49. *For each dominant $\mu \in \mathbb{Z}^n$ there is a unique decomposition*

$$V_\mu = \oplus_{\lambda \in \mathrm{GT}(n)_\mu} V_\lambda$$

of $V_\mu \in \mathrm{Irr}(\mathrm{U}(n))$ into one-dimensional subspaces V_λ, indexed by $\lambda \in \mathrm{GT}(n)_\mu$, with the property that for each $j \in \{1..n\}$ and $\nu \in \mathbb{Z}^{(j)}$, the subspace

$$\oplus_{\lambda \in \mathrm{GT}(n)_\mu : \lambda^{(j)} = \nu} V_\lambda$$

of V_μ is $\mathrm{U}(j)$-invariant and isomorphic as a representation of $\mathrm{U}(j)$ to a direct sum of isomorphic copies of $V_\nu \in \mathrm{Irr}(\mathrm{U}(j))$; it is then the maximal subspace with this property.

Proof. More-or-less immediate from the above discussion; ask me if anything seems unclear. □

This decomposition can be refined further by choosing basis elements for the one-dimensional spaces V_λ and describing the representation explicitly in terms of that basis, but we will not pursue such refinements here.

4. Matrix coefficients for compact groups and the Peter–Weyl theorem

4.1. **Spaces of matrix coefficients.** Let G be a compact group and (π, V) a finite-dimensional representation. So far, our main tool for saying anything interesting about π has been via its character $\chi_\pi : G \to \mathbb{C}$, defined as the composition

$$G \xrightarrow{\pi} \mathrm{GL}(V) \hookrightarrow \mathrm{End}(V) \xrightarrow{\mathrm{trace}} \mathbb{C}.$$

To say more, we replace the trace map by a general linear functional $\alpha \in \mathrm{End}(V)^*$. We denote by $\pi_\alpha : G \to \mathbb{C}$ the resulting composition

$$G \xrightarrow{\pi} \mathrm{GL}(V) \hookrightarrow \mathrm{End}(V) \xrightarrow{\alpha} \mathbb{C}.$$

Definition 50. A *matrix coefficient* of $\pi = (\pi, V)$ is a function $G \to \mathbb{C}$ of the form π_α for some $\alpha \in \mathrm{End}(V)$. We denote by

$$\mathcal{A}(\pi) := \{\pi_\alpha : \alpha \in \mathrm{End}(V)^*\}$$

the space of matrix coefficients of π.

For example, if α is the trace map, then $\pi_\alpha = \chi_\pi$, so the character $\chi_\pi \in \mathcal{A}(\pi)$ is a matrix coefficient.[6]

4.2. **Uniqueness of invariant inner products.** It will be convenient in discussing matrix coefficients to suppose that our representations are unitary, so that we can work simply with orthonormal bases for a given representation rather than bases and dual bases for a representation and its dual. In making this assumption it's convenient to know the following:

Lemma 51. For $(\pi, V) \in \mathrm{Irr}(G)$, any two invariant inner products $\langle, \rangle_1$ and $\langle, \rangle_2$ on V differ by a positive scalar: there exists $C > 0$ so that $\langle u, v \rangle_1 = C \langle u, v \rangle_2$ for all $u, v \in \pi$.

Proof. We may identify inner products $\langle, \rangle$ with certain linear maps

$$V \to \overline{V}^*$$

$$v \mapsto \langle v, \cdot \rangle.$$

The inner product is invariant if and only if the linear map is equivariant, i.e., defines an element of the space $\mathrm{Hom}_G(\pi, \overline{\pi}^*)$. But Schur's lemma implies that the latter space is one-dimensional. Hence any two invariant inner products differ by a scalar. Positive-definiteness forces the scalar to be positive. $\qquad\square$

4.3. **Matrix entries as a basis for the space of matrix coefficients.** We suppose henceforth that π is unitary, and fix an orthonormal basis $\{e_i\}_{i=1}^{\dim(\pi)}$ for V. Each $T \in \mathrm{End}(V)$ then defines a matrix with entries $T_{ij} = \langle Te_j, e_i \rangle$. Write $\varepsilon_{ij} \in \mathrm{End}(V)^*$ for the map $\mathrm{End}(V) \to \mathbb{C}$ assigning to an operator T the matrix entry T_{ij}. Then the ε_{ij} give a basis for $\mathrm{End}(V)^*$, and so the functions $\pi_{ij} := \pi_{\varepsilon_{ij}}$ span the space $\mathcal{A}(\pi)$ of matrix coefficients of π. (They need not in general be linearly independent: consider for instance a direct sum of copies of the trivial representation, for which $\pi_{ii}(g) = 1$ for all i and g.) The numbers $\pi_{ij}(g)$ are the matrix entries of $\pi(g)$. We may write the map trace : $\mathrm{End}(V) \to \mathbb{C}$ as the sum of diagonal matrix entries $\sum_i \varepsilon_{ii}$, so that $\chi_\pi = \sum_i \pi_{ii}$.

---

[6]End of lecture #5, Tuesday, 5 March

4.4. Special functions as matrix coefficients. "Most" interesting special functions in mathematics and mathematical physics (Bessel, Whittaker, Legendre, Laguerre, ...) arise as matrix coefficients of representations, and may be profitably studied from this perspective. For instance, $G = \mathbb{R}/2\pi\mathbb{Z}$ has a representation π on $V = \mathbb{C}^2$ given by the rotations

$$\pi(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \pi_{11}(\theta) & \pi_{12}(\theta) \\ \pi_{21}(\theta) & \pi_{22}(\theta) \end{pmatrix},$$

whose matrix coefficients are thus the trigonometric functions. Their addition law is obtained by writing the homomorphism property $\pi(\theta_1 + \theta_2) = \pi(\theta_1)\pi(\theta_2)$ as $\pi_{ij}(\theta_1 + \theta_2) = \sum_k \pi_{ik}(\theta_1)\pi_{kj}(\theta_2)$, i.e.,

$$\begin{pmatrix} \cos(\theta_1 + \theta_2) & \sin(\theta_1 + \theta_2) \\ -\sin(\theta_1 + \theta_2) & \cos(\theta_1 + \theta_2) \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{pmatrix},$$

giving a convenient way to remember the formulas $\cos(\theta_1 + \theta_2) = \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)$ and $\sin(\theta_1 + \theta_2) = \cos(\theta_1)\sin(\theta_2) + \sin(\theta_1)\cos(\theta_2)$. See the book by Vilenkin for many more examples like this.

4.5. Some actions of $G \times G$. Anyway, the association $\alpha \mapsto \pi_\alpha$ defines a map

$$\operatorname{End}(V)^* \to \mathcal{A}(\pi). \tag{24}$$

The spaces involved in this map are naturally representations of the product group $G \times G$: For $(g_1, g_2) \in G \times G$ and $\alpha \in \operatorname{End}(V)^*$, we denote by $(g_1, g_2) \cdot \alpha$ the functional

$$\operatorname{End}(V) \to \mathbb{C}$$

$$T \mapsto \alpha(\pi(g_1)^{-1} T \pi(g_2))$$

Another way to arrive at the same definition is via the external tensor product $V_1 \boxtimes V_2$ of representations V_1 of G_1 and V_2 of G_2 as in the homework. This is the representation on the tensor product space given by $(g_1, g_2) \cdot (v_1 \otimes v_2) = g_1 v_1 \otimes g_2 v_2$. We have equivariant identifications

$$\operatorname{End}(V)^* \cong (V^* \boxtimes V)^* \cong V^* \boxtimes V. \tag{25}$$

For any function $f : G \to \mathbb{C}$, we write $(g_1, g_2) \cdot f$ for the function

$$G \to \mathbb{C}$$

$$x \mapsto f(g_1^{-1} x g_2).$$

This definition applies to $f \in \mathcal{A}(\pi)$, and the space $\mathcal{A}(\pi)$ is $G \times G$-invariant. The map (24) is equivariant for these actions of $G \times G$.

We denote by ΔG the diagonal subgroup $\{(g, g) : g \in G\}$ of $G \times G$. Then the fixed subspace $L^2(G)^{\Delta G}$ is the space of class functions $L^2(G)^{\text{class}}$.

4.6. Burnside's lemma. The spaces of matrix coefficients are particularly well-behaved in the case of an irreducible representation:

Lemma 52. Let $(\pi, V) \in \operatorname{Irr}(G)$. Then

(i) *$\operatorname{End}(V)^*$ and $\mathcal{A}(\pi)$ define irreducible representations of $G \times G$, and the map $\operatorname{End}(V)^* \to \mathcal{A}(\pi)$ is an isomorphism.*

(ii) *The ΔG-fixed subspace $(\operatorname{End}(V)^*)^{\Delta G}$ is the one-dimensional space $\mathbb{C}$ trace of multiples of $\operatorname{trace} : \operatorname{End}(V) \to \mathbb{C}$, whereas $\mathcal{A}(\pi)^{\Delta G}$ is the one-dimensional space $\mathbb{C}\chi_\pi$ of multiples of the character of π.*

(iii) *(Burnside's lemma) The matrix entries π_{ij} define linearly independent functions on G. The set $\{\pi(g) : g \in G\}$ spans $\mathrm{End}(V)$.*

Proof. (i) From the homework, we know that if V_1 and V_2 are irreducible representations of the (compact) groups G_1 and G_2, then $V_1 \boxtimes V_2$ is an irreducible representation of $G_1 \times G_2$. From this and (25) we deduce that $\mathrm{End}(V)^*$ is $G \times G$-irreducible. The map $\mathrm{End}(V)^* \to \mathcal{A}(\pi)$ is equivariant, surjective (by definition) and nonzero (since its image contains χ_π), hence its kernel W is a proper invariant subspace of $\mathrm{End}(V)^*$, but since $\mathrm{End}(V)^*$ is irreducible, we must have $W = \{0\}$. Therefore the map in question is an isomorphism; in particular, $\mathcal{A}(\pi)$ is irreducible.

(ii) It's clear that trace and χ_π define nonzero elements of $\mathrm{End}(V^*)^{\Delta G}$ and $\mathcal{A}(\pi)^{\Delta G}$, so we just need to check that the latter two spaces are at most one-dimensional. Part (i) implies that they are isomorphic, so we may conclude via the identification $\mathrm{End}(V)^* \cong \mathrm{End}(V^*)$ and Schur's lemma in the form $\dim \mathrm{End}_G(V^*) \leq 1$.

(iii) The π_{ij} are the images of the basis elements ε_{ij} under the isomorphism $\mathrm{End}(V)^* \to \mathcal{A}(\pi)$, so they are linearly independent. For the second assertion, let $W \leq \mathrm{End}(V)$ denote the span of $\{\pi(g) : g \in G\}$. If $W \neq \mathrm{End}(V)$, then we can find a nonzero $\alpha \in \mathrm{End}(V)^*$ vanishing on W. Then $\pi_\alpha(g) = \alpha(\pi(g)) \in \alpha(W) = \{0\}$ for all $g \in G$, so $\pi_\alpha = 0$, contrary to (i). $\qquad\qquad$ $\square$

4.7. Schur orthogonality.

Definition 53. The space $\mathrm{End}(V)^*$ comes with a natural $G \times G$-invariant inner product ("dual of the Hilbert–Schmidt inner product") given explicitly in coordinates by requiring that the ε_{ij} be an orthonormal basis, i.e.,

$$\left\langle \sum a_{ij}\varepsilon_{ij}, \sum b_{ij}\varepsilon_{ij} \right\rangle = \sum a_{ij}\overline{b_{ij}}$$

and more invariantly as the composition

$$\mathrm{End}(V)^* \otimes \overline{\mathrm{End}(V)^*} \cong V \otimes V^* \otimes \overline{V} \otimes \overline{V}^* \to \mathbb{C} \qquad (26)$$

where the second arrow sends $v_1 \otimes v_2 \otimes \overline{v_3} \otimes \overline{v_4}$ to $\langle v_1, v_3 \rangle_V \langle v_2, v_4 \rangle_{V^*}$, where V^* is equipped with the inner product obtained from that on V via the duality isomorphism $V \ni v \mapsto \langle v, \cdot \rangle \in \overline{V}^*$ induced by the given inner product on V.

Another inner product on $\mathrm{End}(V)^*$ is obtained from that on $L^2(G)$ by taking matrix coefficients. It's natural to ask how the two inner products compare:

Theorem 54 (Schur orthogonality relations). *Let $(\pi, V), (\pi', V') \in \mathrm{Irr}(G)$, with G compact as usual. Then for $\alpha \in \mathrm{End}(V)^*$ and $\beta \in \mathrm{End}(V')^*$,*

$$\langle \pi_\alpha, \pi'_\beta \rangle = \begin{cases} 0 & \text{if } \pi \not\cong \pi', \\ \frac{1}{\dim(\pi)}\langle \alpha, \beta \rangle & \text{if } (\pi, V) = (\pi', V'), \end{cases}$$

where the inner product on the left is in $L^2(G)$ with respect to the probability Haar, while $\langle \alpha, \beta \rangle$ denotes the dual Hilbert–Schmidt inner product defined above. Explicitly,

$$\langle \pi_{ij}, \pi'_{k\ell} \rangle = \frac{1}{\dim(\pi)}\delta_{\pi,\pi'}\delta_{i,k}\delta_{j,\ell}. \qquad (27)$$

Proof. The proof is basically as in Lemma 52. If π is not G-isomorphic to π', then $\mathrm{End}(\pi)^*$ is likewise not $G \times G$-isomorphic to $\mathrm{End}(\pi')^*$ (as may be checked for instance by verifying that their characters are orthogonal, as on the homework), but the map

$$\mathrm{End}(V)^* \to \overline{\mathrm{End}(V')^*}^* \cong \mathrm{End}(V')^*$$

$$\alpha \mapsto [\beta \mapsto \langle \pi_\alpha, \pi_\beta \rangle]$$

is $G \times G$-equivariant, so Schur's lemma implies that it vanishes identically. In the case $(\pi, V) = (\pi', V')$, we use Lemma 52 (a basic consequence of Schur's lemma) to write

$$\langle \pi_\alpha, \pi_\beta \rangle = C \langle \alpha, \beta \rangle$$

for some $C > 0$, which may depend upon π, but not upon α and β. To compute C, we take

$$\alpha = \beta = \mathrm{trace},$$

so that

$$\pi_\alpha = \pi_\beta = \chi_\pi,$$

and note that

$$\langle \mathrm{trace}, \mathrm{trace} \rangle = \Big\langle \sum \varepsilon_{ii}, \sum \varepsilon_{ii} \Big\rangle = \dim(\pi)$$

and recall that $\langle \chi_\pi, \chi_\pi \rangle = 1$. $\qquad\square$

4.8. The coefficient ring of a compact group. Having now studied in detail the matrix coefficients of irreducibles, we piece them together as follows.

Definition 55. Let G be, as usual, a compact group. The *coefficient ring* of G is defined to be the set

$$\mathcal{A}(G) := \cup_{(\pi, V) \text{ finite-dimensional}} \mathcal{A}(\pi) \subseteq L^2(G).$$

The terminology is justified by:

Theorem 56. (i) $\mathcal{A}(G)$ *is a $\mathbb{C}$-algebra, closed under complex conjugation, and $G \times G$-invariant.*

(ii) $\mathcal{A}(G)$ *is the orthogonal direct sum in $L^2(G)$ of the subspaces $\mathcal{A}(\pi)$ for $\pi \in \mathrm{Irr}(G)$. The functions*

$$\sqrt{\dim(\pi)}\pi_{ij} : G \to \mathbb{C}, \tag{28}$$

for π running over $\mathrm{Irr}(G)$ and $i, j = 1..\dim(\pi)$, form an orthonormal basis of $\mathcal{A}(G)$. Similarly, $\mathcal{A}(G)^{\Delta G}$ is the orthogonal direct sum in $L^2(G)^{\mathrm{class}}$ of the subspaces $\mathbb{C}\chi_\pi$. In particular, $\mathrm{Irr}(G)$ is at most countably infinite.

(iii) *For any $f \in L^2(G)$, the following are equivalent:*
   (a) $f \in \mathcal{A}(G)$
   (b) f *is* right-finite: $\dim \mathrm{span}\{f(\cdot g) : g \in G\} < \infty.$
   (c) f *is* left-finite: $\dim \mathrm{span}\{f(g^{-1}\cdot) : g \in G\} < \infty.$
   (d) f *is* bi-finite, *or simply* finite: $\dim \mathrm{span}\{f(g_1^{-1} \cdot g_2) : g_1, g_2 \in G\} < \infty.$

Proof. (i) The proof is similar to what we did for characters (Lemma 24): The conjugate of a matrix coefficient is a matrix coefficient for the conjugate representation, namely, for $\alpha \in \mathrm{End}(\pi)^*$, we have $\overline{\pi_\alpha} = \overline{\pi}_{\overline{\alpha}}$ with $\overline{\alpha} \in \overline{\mathrm{End}(\pi)^*} \cong \mathrm{End}(\overline{\pi})^*$, so $\mathcal{A}(G)$ is closed under complex conjugation. Similarly, the sum (resp. product) of two matrix coefficients is a matrix coefficient for the direct sum (resp. product) of the corresponding representations, i.e., $\pi_\alpha + \pi_{\beta'} =$

$(\pi \oplus \pi')_{\alpha \oplus \beta}$ and $\pi_\alpha \pi_{\beta'} = (\pi \otimes \pi')_{\alpha \otimes \beta}$ where $\alpha \oplus \beta \in \operatorname{End}(\pi)^* \oplus \operatorname{End}(\pi')^* \hookrightarrow \operatorname{End}(\pi \oplus \pi')^*$ and $\alpha \otimes \beta \in \operatorname{End}(\pi)^* \otimes \operatorname{End}(\pi')^* \hookrightarrow \operatorname{End}(\pi \otimes \pi')^*$.

(ii) We have seen that finite-dimensional (π, V) may be assumed unitary and then decomposed as orthogonal direct sums $W_1 \oplus \cdots \oplus W_m$ of irreducibles $\sigma_1, \ldots, \sigma_m$. For each $k = 1..m$, fix an orthonormal basis $\{e_i^{(k)}\}_i$ for W_k. Then $\mathcal{A}(\pi)$ is spanned by the functions

$$G \ni g \mapsto \langle e_i^{(k)}, g e_j^{(\ell)} \rangle.$$

We have $g e_j^{(\ell)} \in W_\ell$, which is orthogonal to $e_i^{(k)}$ unless $k = \ell$, in which case the above function belongs to $\mathcal{A}(\sigma_k)$. Thus $\mathcal{A}(\pi) = \mathcal{A}(\sigma_1) + \cdots + \mathcal{A}(\sigma_k)$. Taking unions, we obtain $\mathcal{A}(G) = \sum_{\pi \in \operatorname{Irr}(G)} \mathcal{A}(\pi)$. The required orthogonality follows from the Schur orthogonality relations.

(iii) If $f \in \mathcal{A}(G)$, then $f \in \mathcal{A}(\pi)$ for some finite-dimensional π, so the $G \times G$-span of f is contained in the space $\mathcal{A}(\pi)$ whose dimension is finite (e.g., bounded by $\dim(\pi)^2$). To complete the proof it suffices to verify that if a function f is right- or left-finite, then it belongs to $\mathcal{A}(G)$. We treat the case that f is right-finite, the other case being similar. Let V denote the span of the right translates $f(\cdot g)$ of f, so that $\dim(V) < \infty$ by assumption. Take for π the representation of G on V given by right translation. (Note **as I did not in lecture** that this is actually a representation: V is a finite-dimensional subspace of $L^2(G)$, hence is closed, and so the continuity of the action of G on V follows from that of G on $L^2(G)$.)

We attempt now to define the "act on f, then evaluate at identity" functional $\alpha \in \operatorname{End}(V)^*$ by the formula

$$\alpha(T) := (Tf)(e), \tag{29}$$

with $e \in G$ denoting the identity element. (Note **as I did not in lecture** that this definition does not obviously make sense, because f is merely an "L^2-function," and so is not obviously defined pointwise except off some unspecified set of measure zero. However, as we will verify below (non-circularly) in §5.4.1, our assumptions imply that every element of V is actually represented by a *continuous* function, to which the definition applies.) Then $\pi_\alpha(g) = (\pi(g)f)(e) = f(eg) = f(g)$, so $f = \pi_\alpha \in \mathcal{A}(\pi) \subseteq \mathcal{A}(G)$, as required.

$\square$

4.9. Peter–Weyl theorem: statement and proof sketch. We can now state the more complete form of Theorem 27:

Theorem 57 (Peter–Weyl theorem). *Let G be a compact group.*

(i) *$\mathcal{A}(G)$ is dense in the space $C(G)$ of continuous functions on G equipped with the topology defined by the supremum norm.*

(ii) *$\mathcal{A}(G)$ is dense in $L^2(G)$, hence $L^2(G) = \hat{\oplus}_{\pi \in \operatorname{Irr}(G)} \mathcal{A}(\pi)$; here $\hat{\oplus}$ denotes the "Hilbert direct sum," i.e., the closure of the "ordinary" or "algebraic" direct sum $\oplus$. Thus the functions (28) give an orthonormal basis (in the sense of Hilbert space) for $L^2(G)$.*

(iii) *$\mathcal{A}(G)^{\Delta G} = \oplus_{\pi \in \operatorname{Irr}(G)} \mathbb{C} \chi_\pi$ is dense in $L^2(G)^{\operatorname{class}}$ (thus $L^2(G)^{\operatorname{class}} = \hat{\oplus}_{\pi \in \operatorname{Irr}(G)} \mathbb{C} \chi_\pi$) and also dense in $C(G)^{\operatorname{class}}$.*

(iv) If (π, V) is any irreducible representation of G on any space V such that "one has a reasonable theory of V-valued integrals" (e.g., Hilbert, Banach, Frechet, or "locally convex quasi-complete"), then $\dim(V) < \infty$.

The main assertion here is that $\mathcal{A}(G)$ is dense in $L^2(G)$; we will deduce the remaining assertions either from this one or via easy modifications of its proof. As "warm-up," we observe some easy special cases:

- If G is a finite group, then both $\mathcal{A}(G)$ and $L^2(G)$ consist of all functions $G \to \mathbb{C}$, so there is nothing to show.
- A *profinite group* G is a compact group which admits an open basis $\mathcal{U}$ at the identity element consisting of compact open normal subgroups U. (This might not be the standard definition, but is equivalent to it, and suits the purposes of our discussion.) For instance, the group $G = \mathrm{GL}_n(\mathbb{Z}_p)$, which may be identified with the inverse limit of the groups $\Gamma_k = \mathrm{GL}_n(\mathbb{Z}/p^k\mathbb{Z})$ as k runs over the positive integers, is profinite, with the compact open normal subgroups $U_k := \ker(G \to \Gamma_k)$ giving a neighborhood basis of the identity.

  For any profinite G and any $U \in \mathcal{U}$, the quotient group G/U is finite; indeed, G is compact and $G = \cup_{g \in G/U} gU$ is an open cover. Because continuous functions on a compact space are uniformly continuous, we may find for each $f \in C(G)$ and $\varepsilon > 0$ a subgroup $U \in \mathcal{U}$ so that f varies by at most ε on each coset of U, i.e.,

$$|f(gu) - f(g)| \leq \varepsilon \text{ for all } g \in G \text{ and } u \in U.$$

  Let $C(G)^U \subseteq C(G)$ denote the space of functions $G \to \mathbb{C}$ that are constant on the cosets of U, and let $f_U \in C(G)^U$ be defined by averaging f over U-cosets, i.e.,

$$f_U(g) := \mathbb{E}_{u \in U} f(gu),$$

  where $\mathbb{E}$ denotes the integral with respect to the probability Haar on the compact group U. Then $\sup |f - f_U| \leq \varepsilon$. Thus $\cup_{U \in \mathcal{U}} C(G)^U$ is dense in $C(G)$. On the other hand, the spaces $C(G)^U$ are finite-dimensional and $G \times G$-invariant, hence contained in $\mathcal{A}(G)$. It follows that $\mathcal{A}(G)$ is dense in $C(G)$.

The proof for general compact groups G will be similar to what we just did in the profinite case, but a bit more technically involved because we don't in general have such a spectacularly convenient neighborhood basis of subgroups. (Think of U(1), which visibly has no nontrivial subgroups contained in $\{z \in \mathrm{U}(1) : |z-1| < 1/10\}$.) We'll construct instead some "approximate analogues" of the maps $f \mapsto f_U$ by convolving f with functions $\phi \in C(G)$ that are supported in small neighborhoods of the identity element and satisfy the normalization $\int_G \phi = 1$. Set $T_\phi f(g) := \int_{g \in G} \phi(g) f(g)\, dg$, where dg denotes the probability Haar on G. (For instance, if G is profinite and $\phi = \mathrm{vol}(U, dg)^{-1} 1_U$ for some $U \in \mathcal{U}$, then $T_\phi f = f_U$ as above.) If ϕ is sufficiently concentrated, then $T_\phi f$ will approximate f. Since G is compact, we may assume that ϕ is a class function (corresponding to U being normal). We might as well assume also that ϕ is real-valued and invariant by inversion (corresponding to U being closed under inversion). We'll see then that T_ϕ defines a compact self-adjoint equivariant operator on $L^2(G)$. The spectral theory of such operators will then allow us to approximate $T_\phi f$ by its projection to the eigenspaces of T_ϕ of eigenvalue λ

with $|\lambda| > \varepsilon$; since the eigenspaces of such operators with nonzero eigenvalue are finite-dimensional and G-invariant, this gives the required approximation of $T_\phi f$ and hence also of f via finite functions. Details next time.[7]

Let's note one final easy special case. Suppose that G is a compact *matrix* group, i.e., a compact subgroup of some $\mathrm{GL}_n(\mathbb{C})$. (By Theorem 10 applied to the identity representation $G \to \mathrm{GL}_n(\mathbb{C})$, we might assume further that G is contained in the unitary group $\mathrm{U}(n)$.) Let $\mathcal{O}$ denote the space of functions $f : G \to \mathbb{C}$ given by the restriction $f = \tilde{f}|_G$ of some function $\tilde{f} : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}$ for which $\tilde{f}(g)$ is a polynomial function of the matrix entries of g. We may write $\mathcal{O} = \cup_{d \geq 0} \mathcal{O}_d$, where $\mathcal{O}_d$ denotes the subspace of polynomials of degree at most d. Each space $\mathcal{O}_d$ is finite-dimensional and G-invariant, hence $\mathcal{O}_d \subseteq \mathcal{A}(G)$, hence $\mathcal{O} \subseteq \mathcal{A}(G)$. Since $\mathcal{O}$ is a $\mathbb{C}$-algebra that is closed under complex conjugation and separates points on G, we deduce by Stone–Weierstrass that $\mathcal{O}$ is dense in $C(G)$. It follows also that $\mathcal{A}(G)$ is dense in $C(G)$ (and that $\mathcal{A}(G) = \mathcal{O}$).

On a related note, we observe the following consequence of Peter–Weyl:

Corollary 58. *Let G be a compact Lie group. Then G is isomorphic to a compact matrix Lie group.*

Thus Peter–Weyl for compact *Lie* groups G is essentially equivalent to the fact that any such G is isomorphic to a matrix Lie group.

Proof. Enumerate $\mathrm{Irr}(G)$ as $V_1, V_2, \ldots$, and let $G_n \leq G$ denote the kernel of the map $G \to \mathrm{GL}(V_1 \oplus \cdots \oplus V_n)$ given by the direct sum of the representations $V_1, \ldots, V_n$. Then G_n is a closed subgroup of G, hence a Lie subgroup, so we may speak of its Lie algebra $\mathfrak{g}_n$, which is a subalgebra of the Lie algebra $\mathfrak{g}$ of G. The subgroups G_n decrease with n, hence so do the subspaces $\mathfrak{g}_n$. By Peter–Weyl, that the matrix coefficients of the irreducible representations separate points in G. It follows that $\cap_n G_n = \{1\}$ and thus $\cap_n \mathfrak{g}_n = \{0\}$. Since the $\mathfrak{g}_n$ decrease and are finite-dimensional, we may find n_0 so that $\mathfrak{g}_n = \{0\}$ for $n \geq n_0$; for such n, the Lie group G_n is thus compact and zero-dimensional, hence finite. Using again that $\cap_n G_n = \{1\}$, we may find n so that $G_n = \{1\}$. Then $G \to \mathrm{GL}(V_1 \oplus \cdots \oplus V_n)$ is an injective homomorphism with compact domain (and Hausdorff codomain), hence defines an isomorphism (of topological groups) onto its image. $\qquad\square$

The proof shows also that any compact group G is isomorphic to an inverse limit of compact matrix Lie groups.

We'll give the proof of Peter–Weyl below after developing some preliminaries concerning integral operators, which are of independent interest.

---

[7]End of lecture #6, Thursday, 7 March

5. Integral operators

5.1. **Integrating functions taking values in Hilbert spaces.**

Lemma 59. *Let (Ω, μ) be a locally compact space equipped with a Radon measure, let V be a Hilbert space, and let f be an element of the space $C_c(\Omega \to V)$ of compactly-supported continuous functions from Ω to V. Then there is a unique element $v \in V$ so that for each element ℓ of the continuous dual V^*, we have*

$$\ell(v) = \int_{x \in \Omega} \ell(f(x)) \, d\mu(x). \qquad (30)$$

We then use the notation $\int_\Omega f \, d\mu := v \in V$.

Proof. We may define v first as an algebraic functional $V^* \to \mathbb{C}$ via (30). By Hilbert space duality, our task reduces to showing that v belongs to the double dual of V, i.e., that $|\ell(v)| \le C\|\ell\|$ for some constant $C \ge 0$ and all $\ell \in V^*$. We may take $C = C_1 C_2$, where $C_1 := \mu(\mathrm{supp}(f)) < \infty$ and $C_2 := \max_{x \in \mathrm{supp}(f)} \|f(x)\| < \infty$. $\square$

This lemma and hence our subsequent development can be generalized further (to Banach spaces, Frechet spaces, etc.), but the present generality should suffice for our aims.

5.2. **Definition of integral operators attached to representations.** From now on we take for G a locally compact group and (π, V) a *Hilbert representation*; by this we mean that V is a Hilbert space, without requiring that π be unitary.

Lemma 60. *For every compactly-supported signed Radon measure μ on G there is a unique bounded linear operator $\pi(\mu) : V \to V$ such that*

$$\ell(\pi(\mu)v) = \int_{g \in G} \ell(\pi(g)v) \, d\mu(g)$$

for all $v \in V$ and $\ell \in V$. We then use the notation $\pi(\mu) =: \int_{g \in G} \pi(g) \, d\mu(g)$.

Proof. Set $\Omega := \mathrm{supp}(\mu)$. For each $v \in V$, we may apply the previous lemma to the function $f_v : \Omega \to V$ given by $g \mapsto \pi(g)v$. We obtain elements $\pi(\mu)v \in V$ and a linear map $\pi(\mu) : V \to V$. It remains to show that $\pi(\mu)$ is bounded. For this it suffices to show that there exists $C \ge 0$ (depending upon Ω) so that the operator norm $\|\pi(g)\|$ is bounded by C for all $g \in \Omega$. The conclusion follows by Banach–Steinhaus (presumably covered in the functional analysis course), which we record here for convenience:

A pointwise bounded family $(T_\alpha)_{\alpha \in I}$ of bounded linear maps $T_\alpha : X \to Y$ from a Banach space X to a normed space Y is uniformly bounded, i.e., if for each $x \in X$ there exists $C(x) \ge 0$ so that $\|T_\alpha x\| \le C(x)\|x\|$ for all α, then there exists $C \ge 0$ so that $\|T_\alpha x\| \le C\|x\|$ for all α, x.

We apply this to the family $(\pi(g))_{g \in \Omega}$ of bounded linear maps $\pi(g) : V \to V$; the pointwise boundedness property follows from the continuity of $G \times V \to V$, while the uniform boundedness gives the conclusion that we seek.

For convenience and as a reminder of what's going on "under the hood," we record a proof of Banach–Steinhaus as quoted above. (This was not presented in lecture, and is not otherwise relevant to the course.) If the conclusion fails, then we can find a sequence of nonzero vectors $x_n \in X$ and elements T_n of the family

(T_α) so that $\|T_n x_n\|/\|x_n\| \to \infty$. After normalizing the x_n by a suitable scalar, we may arrange that

$$\|x_n\| \to 0, \quad \|T_n x_n\| \to \infty.$$

We now inductively choose a "sufficiently sparse" subsequence $(x_{n_k})_{k \geq 1}$, as follows. For $k \geq 1$, having chosen $n_1 < \cdots < n_{k-1}$, choose n_k sufficiently large that

- $\|x_{n_k}\| \leq 2^{-k}$,
- $\sum_{\ell=1..k-1} \|T_{n_\ell} x_{n_k}\| \leq 2^{-k}$ (as we may, because each T_{n_ℓ} is bounded), and
- $\|T_{n_k} x_{n_k}\| \geq \sum_{\ell=1..k-1} \|T_{n_k} x_{n_\ell}\| + 2^k$ (as we may, because the family (T_α) is assumed pointwise bounded).

Since $\sum_k \|x_{n_k}\| < \infty$ and X is complete, we have $x := \sum_k x_{n_k} \in X$. But

$$\|T_{n_k} x\| \geq \|T_{n_k} x_{n_k}\| - \sum_{\ell=1..k-1} \|T_{n_k} x_{n_\ell}\| - \sum_{\ell \geq k+1} \|T_{n_k} x_{n_\ell}\| \geq 2^k - \sum_{\ell \geq k+1} 2^{-\ell} \to \infty,$$

contrary to the assumed pointwise boundedness of the family (T_α). $\qquad\square$

5.3. Basic properties.
From now on we assume moreover (mainly for convenience) that G is *unimodular*, so that we may fix a (left and right) Haar measure dg. For $f \in C_c(G)$, the above discussion then applies to $\mu = f\, dg$, giving us operators on V that we denote by $\pi(f) := \pi(f\, dg)$, thus $\pi(f) = \int_{g \in G} f(g)\pi(g)\, dg$.

Lemma 61. *Let* $f, f_1, f_2 \in C_c(G)$ *and* $g \in G$.

(i) $\pi(f_1)\pi(f_2) = \pi(f_1 * f_2)$, *where* $f_1 * f_2 \in C_c(G)$ *denotes the convolution*

$$f_1 * f_2(x) := \int_{g \in G} f_1(g) f_2(g^{-1}x)\, dg.$$

(ii) $\pi(g)\pi(f) = \pi(f(g^{-1}\cdot))$, $\pi(f)\pi(g) = \pi(f(\cdot g^{-1}))$.
(iii) *If* $f \in C_c(G)^{\text{class}}$, *then* $\pi(f)$ *is equivariant.*
(iv) *If* (π, V) *is unitary, then the adjoint* $\pi(f)^*$ *is given by* $\pi(f^*)$, *where* $f^*(x) := \overline{f(x^{-1})}$.

Proof. (i) By substituting $g_2 \mapsto g_1^{-1} g_2$,

$$\pi(f_1)\pi(f_2) = \int\int f_1(g_2) f_2(g_2) \underbrace{\pi(g_1)\pi(g_2)}\, \pi(g_1 g_2)\, dg_1\, dg_2$$

$$= \int (\underbrace{\int f_1(g_1) f_2(g_1^{-1} g_2)\, dg_1}_{(f_1 * f_2)(g_2)})\pi(g_2)\, dg_2$$

$$= \pi(f_1 * f_2).$$

(ii) By substituting $x \mapsto g^{-1}x$, $\pi(g)\pi(f) = \int f(x)\pi(g)\pi(x)\, dx = \int f(g^{-1}x)\pi(x)\, dx = \pi(f(g^{-1}\cdot))$; similarly for the other assertion.
(iii) We've seen that $\pi(g)\pi(f)\pi(g^{-1}) = \pi([x \mapsto f(g^{-1}xg)])$, which coincides with $\pi(f)$ if f is a class function.
(iv) Since π is unitary, we have $\pi(g)^* = \pi(g^{-1})$, hence $\pi(f)^* = \int \overline{f(g)}\pi(g^{-1})\, dg$; by the change of variables $g \mapsto g^{-1}$, this coincides with $\pi(f^*)$.

$\qquad\square$

5.4. Approximating vectors by their images under integral operators.

Lemma 62. *For each $v \in V$ and $\varepsilon > 0$ there exists $f \in C_c(G)$ so that $\|\pi(f)v - v\| \leq \varepsilon$*

Proof. By the continuity of the action, we may find a neighborhood U of the identity element in G so that $\|\pi(g)v - v\| \leq \varepsilon$ for all $g \in U$. By Urysohn, we may find $f \in C_c(U)$ such that $\int_G f \, dg = 1$ and $f \geq 0$. Then

$$\|\pi(f)v - v\| = \|\int_{g \in G} f(g)(\pi(g)v - v) \, dg\| \leq (\int f \, dg)\varepsilon = \varepsilon.$$

$\square$

Lemma 63. *In lemma (62), we may arrange moreover that $f \geq 0$ and $f(g) = f(g^{-1})$, and for G compact also that $f \in C_c(G)^{\text{class}}$.*

Proof. Exercise. $\square$

These results may be generalized and refined further. For instance, if G is a Lie group, then we may assume that f is smooth.

5.4.1. *Automatic continuity for finite-dimensional subrepresentations of $L^2(G)$.* We now fulfill the promise made after the statement of Theorem 56 by explaining why for a compact group G, the finiteness (right, left or bi) of $f \in L^2(G)$ implies that f is represented by a continuous function. Suppose for instance that f is right-finite, so that it is contained in some finite-dimensional subspace $V \subseteq L^2(G)$ invariant by the right regular representation ρ of G. For each $v \in V$ and $\varepsilon > 0$, we see by Lemma 62 that there exists $\phi \in C(G)$ so that $\|\rho(\phi)v - v\|_{L^2} \leq \varepsilon$. In particular, the subspace of V consisting of elements of the form $\rho(\phi)v$ is dense. Since V is finite-dimensional, it follows that every element of V is of this form. Note that each such element is pointwise defined.

We verify now that every element of V, say $v' = \rho(\phi)v$, is in fact continuous. For $g \in G$, we have (by Lemma 61) $\rho(g)v' = \rho(\phi(g^{-1}\cdot))v$, hence

$$\sup_{x \in G} |v'(xg) - v'(x)| = \|\rho(g)v' - v'\|_{L^\infty} = \|\rho(\phi(g^{-1}\cdot) - \phi)v\|_{L^\infty} \leq \|\rho(\phi(g^{-1}\cdot) - \phi)\|_{L^2}\|v\|_{L^2},$$

in the last step by Cauchy–Schwarz. Since the left regular representation of G on $L^2(G)$ is continuous, the required continuity of v' follows.

5.5. Some functional-analytic considerations.
Recall (from, e.g., §1.5) the definition of a compact operator on a Hilbert space.

I don't think the following definition is standard, but I like it.

Definition 64. We say that (π, V) is of *compact type* if $\pi(f)$ is compact for each $f \in C_c(G)$.

We recall a bit more functional analysis background.

Lemma 65. *Let (X, μ) be a locally compact space equipped with a positive Borel measure μ such that $V := L^2(X)$ is a separable Hilbert space. Let $k \in L^2(X \times X)$. Then we may define a bounded operator $T : V \to V$ by $Tv(x) = \int_y k(x, y) \, d\mu(y)$. Any such operator is compact.*

Proof sketch; omitted in lecture. Choose an orthonormal basis $e_1, e_2, \ldots$ for $L^2(X)$ and set $a_{ij} := \langle Te_i, e_j \rangle$. Then $\sum_{i,j} |a_{ij}|^2 = \|k\|^2_{L^2(X \times X)} < \infty$. Let $v = \sum_j v_j e_j \in V$. Then $Tv = \sum_i b_i e_i$ with $b_i := \sum_j a_{ij} v_j$. By Cauchy–Schwartz, $|b_i|^2 \leq \|v\|^2 C_i$, where $C_i := \sum_j |a_{ij}|^2$. Thus the image under T of the unit ball in V is contained in the set $\{\sum b_i e_i : |b_i|^2 \leq C_i\}$. Using that $\sum_i C_i < \infty$ and a diagonalization argument, we verify readily that this set is precompact. □

Lemma 66. *If G is a compact group, then its right regular representation $(\rho, L^2(G))$ is of compact type.*

Proof. For $f \in C_c(G) = C(G)$ and $v \in L^2(G)$, we may write $\rho(f)v(x) = \int f(y)v(xy)\, dy = \int v(y)k(x,y)\, dy$, where $k(x,y) := f(x^{-1}y)$. Thus the previous lemma gives the required conclusion. (Sketch of a slightly more direct proof, omitted in lecture: since G is compact and k is continuous, it is uniformly continuous, and so the function $(x,y) \mapsto k(x^{-1}y)$ may be approximated uniformly by step functions; thus $\rho(f)$ is approximated in norm by finite rank operators, hence is compact.) □

5.6. Proof of the L^2-density part of the Peter–Weyl theorem.

We now prove part (ii) of Theorem 57. Let G be a compact group, $v \in L^2(G)$ and $\varepsilon > 0$. Let ρ denote the right regular representation. By lemma 63, we may find $f \in C_c(G)^{\text{class}}$ so that $f \geq 0, f(g^{-1}) = f(g)$ and

$$\|\rho(f)v - v\| \leq \varepsilon. \tag{31}$$

Then $f^* = f$, so $T := \rho(f)$ is self-adjoint. By lemma 66, T is compact. Set $V := L^2(G)$. By Theorem 4, we may write V as the orthogonal Hilbert direct sum of the eigenspaces V_λ of T, taken over $\lambda \in \mathbb{R}$; moreover, setting $V' := \oplus_{|\lambda| \geq \varepsilon} V_\lambda$ and $V'' := \oplus_{|\lambda| < \varepsilon} V_\lambda$, the space V' is finite-dimensional. Since f is a class function, we know (by lemma 61) that the operator T is equivariant (i.e., $\rho(g)T = T\rho(g)$ for all $g \in G$), thus its eigenspaces V_λ and thus their sums V' and V'' are $\rho(G)$-invariant. Thus

$$V' \subseteq \mathcal{A}(G). \tag{32}$$

Let $v' \in V'$ and $v'' \in V''$ denote the components of v, so that $v = v' + v''$. We have

$$\|Tv''\| = \|\sum_{|\lambda| < \varepsilon} \lambda v_\lambda\| \leq \varepsilon\|v\|. \tag{33}$$

By (31) and (33), we deduce that

$$\|v - Tv'\| \leq \varepsilon(1 + \|v\|).$$

On the other hand, $Tv' = \sum_{|\lambda| \geq \varepsilon} \lambda v_\lambda \in V' \subseteq \mathcal{A}(G)$. Thus $\mathcal{A}(G)$ is dense in $L^2(G)$.[8]

5.7. Proof of the uniform density part of the Peter–Weyl theorem.

We now deduce part (i) of Theorem 57. We retain the above notation. Let $v \in C(G)$ and $\varepsilon > 0$. Since G is compact, v is uniformly continuous; by an easy variant of Lemma 62, we may thus find $f \in C(G)$ so that $\|\rho(f)v - v\|_{L^\infty} \leq \varepsilon$. By the L^2-density part of the Peter–Weyl theorem that we have already proved, we may find $f' \in \mathcal{A}(G)$ so that $\|f - f'\|_{L^2} \leq \varepsilon$. Since (G, dg) is a probability space, it follows in particular that $\|f - f'\|_{L^1} \leq \varepsilon$. Since

$$\rho(f)v - \rho(f')v = \int_{f \in G} (f - f')(g)\pi(g)v\, dg$$

_____

[8] End of lecture #7, Tuesday, 12 March

and $\|\pi(g)v\|_{L^\infty} = \|v\|_{L^\infty}$, it follows that

$$\|\rho(f)v - \rho(f')v\|_{L^\infty} \le \|f - f'\|_{L^1}\|v\|_{L^\infty} \le \varepsilon\|v\|_{L^\infty}.$$

Thus

$$\|\rho(f')v - v\| \le \varepsilon(1 + \|v\|_{L^\infty}).$$

Observe finally that for any $g \in G$, we have

$$\rho(g)\rho(f')v = \rho(f'(g^{-1}\cdot))v.$$

Thus if W is any left-G-invariant subspace of $\mathcal{A}(G)$ that contains f', then $\rho(f')v$ belongs to the $\rho(G)$-invariant finite-dimensional space $\rho(W)v$, and so $\rho(f')v \in \mathcal{A}(G)$. We conclude as required that $\mathcal{A}(G)$ is dense in $C(G)$.

Part (iii) of the Peter–Weyl theorem can be proved analogously, or deduced from parts (i) and (ii) by a simple averaging trick; we'll take the latter approach. Given an element v of $L^2(G)^{\mathrm{class}}$ or $C(G)^{\mathrm{class}}$ and $\varepsilon > 0$, we first find $v' \in \mathcal{A}(G)$ so that $\|v' - v\| \le \varepsilon$, where $\|.\|$ denotes either the L^2-norm or L^∞-norm. We then introduce the averaging operator

$$\mathbb{E} : \{\text{functions } f : G \to \mathbb{C}\} \to \{\text{class functions } \mathbb{E}f : G \to \mathbb{C}\}$$

$$\mathbb{E}f(x) := \int_{g \in G} f(gxg^{-1})\, dg.$$

Since f and $x \mapsto f(gxg^{-1})$ have the same norms, the triangle inequality implies that $\|\mathbb{E}f\| \le \|f\|$. Set $v'' := \mathbb{E}v'$. Then $v'' \in \mathcal{A}(G)^{\mathrm{class}}$. Since v is a class function, we have $v = \mathbb{E}v$, and so $(v'' - v) = \mathbb{E}(v' - v)$, hence $\|v'' - v\| \le \|v' - v\| \le \varepsilon$, giving the required approximation of v by an element of $\mathcal{A}(G)^{\mathrm{class}}$.

5.8. **Finite-dimensionality of irreducibles.** We prove part (iv) of Peter–Weyl in the special case of an irreducible Hilbert representation (π, V); the more general conclusion can be deduced similarly after developing integration more generally. The proof is very similar to that of the $C(G)$-density assertion (indeed, a unified statement and proof involving Banach space representations could be given, but I feel like it doesn't hurt to see the argument essentially repeated, since the method is so widely applicable). Let v be a nonzero element of V, and let $\varepsilon > 0$ be sufficiently small. We can find an open neighborhood U of the identity element of G so that $\|\pi(g)v - v\| \le \varepsilon$ for all $g \in U$. Take $f \in C_c(U)$ with $f \ge 0, \int f = 1$. Then $\|\pi(f)v - v\| \le \varepsilon$. Use the L^2-density part of the Peter–Weyl theorem to produce $f' \in C(G)$ so that $\|f - f'\|_{L^2} \le \varepsilon$. Then likewise $\|f - f'\|_{L^1} \le \varepsilon$, and so, as before, $\|\pi(f')v - v\| \le \varepsilon(1 + \|v\|)$. For small enough ε, this implies in particular that $\pi(f')v \ne 0$. Since f' belongs to $\mathcal{A}(G)$, it is contained in some finite-dimensional $G \times G$-invariant subspace W. Then $\pi(f')v$ belongs to the space $\pi(W)v := \{\pi(w)v : w \in W\}$. The latter space is finite-dimensional (with dimension bounded by $\dim(W)$) and nonzero (because it contains $\pi(f')v$). Most significantly, it is invariant: for $w \in W$ and $g \in G$, we have (by Lemma 61)

$$\pi(g)\pi(w)v = \pi(w(g^{-1}\cdot))v.$$

Since V is irreducible, this forces $V = \rho(W)v$, hence $\dim(V) \le \dim(W) < \infty$, as required.

5.9. **Fourier analysis on a compact group** G. Given $f \in C(G)$ (say), its Fourier transform is defined to be the collection $(\pi(f))_{\pi \in \mathrm{Irr}(G)}$ of integral operators that it induces on the (equivalence classes of) finite-dimensional irreducible representations of G. We can restate some of our results concerning orthogonality and completeness of matrix coefficients in terms of these collections of operators. In preparation for doing so, we define on any finite-dimensional Hilbert space V the *Hilbert–Schmidt inner product* of $T_1, T_2 \in \mathrm{End}(V)$ by the formula

$$\langle T_1, T_2 \rangle := \mathrm{trace}(T_1 T_2^*),$$

where as before T_2^* denotes the hermitian adjoint, given with respect to an orthonormal basis by the conjugate transpose. If T_1 and T_2 are respected by matrices (a_{ij}) and (b_{ij}) defined with respect to an orthonormal basis, then we may verify readily that

$$\langle T_1, T_2 \rangle = \sum_{i,j} a_{ij} \overline{b_{ij}}.$$

This is the dual of the inner product considered in §4.7.

Theorem 67. *(i) For $f_1, f_2 \in C(G)$,*

$$\langle f_1, f_2 \rangle_{L^2(G)} = \sum_{\pi \in \mathrm{Irr}(G)} \dim(\pi) \langle \pi(f_1), \pi(f_2) \rangle.$$

(ii) Under "some assumptions" on f,

$$f(1) = \sum_{\pi \in \mathrm{Irr}(G)} \dim(\pi) \chi_\pi(f), \tag{34}$$

where

$$\chi_\pi(f) := \mathrm{trace}(\pi(f)) = \int_{g \in G} f(g) \chi_\pi(g) \, dg.$$

For instance, a sufficient assumption is that f have the form $f_1 * f_2$ for some $f_1, f_2 \in L^2(G)$.

(iii) For $\sigma \in \mathrm{Irr}(G)$, define $\alpha_\sigma := \dim(\sigma)\overline{\chi_\sigma} \in C(G)$. Then for each $\pi \in \mathrm{Irr}(G)$, we have

$$\pi(\alpha_\sigma) = \begin{cases} 0 & \text{if } \pi \not\cong \sigma, \\ 1 & \text{if } \pi \cong \sigma. \end{cases} \tag{35}$$

Moreover, for $\sigma_1, \sigma_2 \in \mathrm{Irr}(G)$,

$$\alpha_{\sigma_1} * \alpha_{\sigma_2} = \begin{cases} \alpha_\sigma & \text{if } \sigma_1 = \sigma_2 =: \sigma, \\ 0 & \text{otherwise.} \end{cases} \tag{36}$$

Proof. (i) With respect to an orthonormal basis of π, we have

$$\pi(f)_{ij} = \int_{g \in G} f(g) \pi_{ij}(g) \, dg,$$

so $\langle \pi(f_1), \pi(f_2) \rangle = \langle f_1, \overline{\pi_{ij}} \rangle \langle \overline{\pi_{ij}}, f_2 \rangle$. The conclusion follows from part (i) of the Peter–Weyl theorem, which we have seen implies that the function $\sqrt{\dim(\pi)}\pi_{ij}$ (and hence likewise the functions $\sqrt{\dim(\pi)}\overline{\pi_{ij}}$) form an orthonormal basis of $L^2(G)$.

(ii) Suppose for instance that $f = f_1 * f_2$ with $f_1, f_2 \in L^2(G)$. Then $f(1) = \int_{g \in G} f_1(g) f_2(g^{-1}) \, dg = \langle f_1, f_2^* \rangle$, while $\langle \pi(f_1), \pi(f_2^*) \rangle = \mathrm{trace}(\pi(f_1)\pi(f_2)) = \mathrm{trace}(\pi(f_1 * f_2)) = \chi_\pi(f)$; the required conclusion thus follows from part (i).

(iii) Using the expansion $\chi_\sigma = \sum_k \sigma_{kk}$, we compute that

$$
\begin{aligned}
\pi(\alpha_\sigma)_{ij} &= \int_{g \in G} \alpha_\sigma(g) \pi_{ij}(g) \, dg \\
&= \sum_k \dim(\sigma) \langle \pi_{ij}, \sigma_{kk} \rangle \\
&= \delta_{\sigma,\pi} \sum_k \delta_{ik} \delta_{jk} \\
&= \delta_{\sigma,\pi} \delta_{ij}.
\end{aligned}
$$

These matrix entries are of the required form.

To prove (36), it suffices by the injectivity of the association $f \mapsto (\pi(f))_{\pi \in \mathrm{Irr}(G)}$ to check that $\pi(\alpha_{\sigma_1} * \alpha_{\sigma_2})$ is given by $\pi(\alpha_\sigma)$ when $\sigma_1 \cong \sigma_2$ and vanishes otherwise, which follows from (35). $\qquad\square$

The pointwise decomposition (34) fails in general for $f \in C(G)$, as one can see already when $G = \mathrm{U}(1)$; there are well-known constructions of continuous functions on the circle whose Fourier series do not converge pointwise. A sufficient condition for the validity of (34) is that f be smooth. (A *smooth function* on a compact Lie group is what you'd expect. A general compact group G can be written, as indicated in §4.9, as an inverse limit of compact Lie groups G_n; a smooth function on G is then a function that factors through a smooth function on one of the quotients G_n of G.) The proof isn't so different from the case $G = \mathrm{U}(1)$, but we won't go into it. Lipschitz continuity should also suffice.

5.10. Isotypic decomposition.

Definition 68. Let (π, V) be a representation of a compact group G, and let $\sigma \in \mathrm{Irr}(G)$. We say that a nonzero vector $v \in V$ is σ-*isotypic* if there are closed invariant subspaces $W_1, \dots, W_n$ of V, with each W_j isomorphic to σ, so that $v \in \sum_j W_j$. We write $V(\sigma) := \{0\} \cup \{\sigma\text{-isotypic } v \in V\}$.

Theorem 69. *Let (π, V) be a Hilbert representation of the compact group G.*

(i) *$\pi(\alpha_\sigma)$ defines a projection $V \to V(\sigma)$; if π is unitary, then it is the orthogonal projection.*

(ii) *$V = \hat{\oplus}_{\sigma \in \mathrm{Irr}(G)} V(\sigma)$.*

Proof. (i) If $v \in V$, then $\pi(\alpha_\sigma)v$ is contained in the space $\pi(\mathcal{A}(\bar\sigma))v$. By Lemma 61), the map

$$
\mathcal{A}(\bar\sigma) \ni f \mapsto \pi(f)v \in V
$$

is equivariant for the left regular representation on the domain. We have seen that $\mathcal{A}(\bar\sigma) \cong \mathcal{A}(\sigma^*) \cong \mathrm{End}(\sigma^*) \cong \sigma \otimes \sigma^*$ as $G \times G$-representations. Under the left regular representation, $\mathcal{A}(\bar\sigma)$ is thus isomorphic to $\sigma^{\oplus \dim(\sigma)}$, so we may write $v = \sum_{j=1}^{\dim(\sigma)} \pi(f_j)v$ with each f_j in a subspace $\tilde{W}_j$ of $\mathcal{A}(\bar\sigma)$ isomorphic to σ. Setting $v_j := \pi(f_j)v \in W_j := \pi(\tilde{W}_j)$, we have $v = \sum v_j$. If W_j is nonzero, then the irreducibility of σ implies that $W_j \cong \sigma$. It follows that v is σ-isotypic.

If $v \in V(\sigma)$, then $\pi(\alpha_\sigma)v = v$, because $\pi(\alpha_\sigma)$ acts by the identity on each of the subspaces W_j that arise in the definition of "σ-isotypic."

We have shown that $\pi(\alpha_\sigma)$ defines a projection $V \to V(\sigma)$. We have $\alpha_\sigma^*(g) = \overline{\alpha_\sigma(g^{-1})} = \dim(\sigma)\overline{\chi_\sigma(g^{-1})} = \dim(\sigma)\overline{\chi_\sigma(g)} = \alpha_\sigma(g)$, so if π is unitary, then (by Lemma 61) $\pi(\alpha_\sigma)$ is self-adjoint, and so defines the orthogonal projection.

(ii) We argue essentially as in §5.8: for each $v \in V$ and $\varepsilon > 0$, we can find $f \in C(G)^{\text{class}}$ so that $\|\pi(f)v - v\| \leq \varepsilon$. We can then decompose f in $L^2(G)^{\text{class}}$ as $\sum_{\sigma \in \text{Irr}(G)} c_\sigma \alpha_\sigma$, say; in particular, we can find a finite subset Σ of $\text{Irr}(G)$ so that $f' := \sum_{\sigma \in \Sigma} c_\sigma \alpha_\sigma$ satisfies $\|f' - f\|_{L^2} \leq \varepsilon$. Arguing as before, we get $\|\pi(f')v - v\| \leq \varepsilon(1 + \|v\|)$. But $\pi(f')v = \sum_{\sigma \in \Sigma} c_\sigma \pi(\alpha_\sigma)v \in \oplus_{\sigma \in \Sigma} V(\sigma)$. Thus $\sum V(\sigma)$ is dense in V. The sum is direct because if $v \in V(\sigma_1) \cap V(\sigma_2)$ with $\sigma_1 \not\cong \sigma_2$, then $v = \pi(\alpha_{\sigma_1})\pi(\alpha_{\sigma_2})v = \pi(\alpha_{\sigma_1} * \alpha_{\sigma_2})v$, which vanishes thanks to (36). $\square$

[9]

5.11. **Decompositions of compact-type representations.** The decompositions of representations of a compact group noted above (e.g., in theorems 22, 57 and 69) do not hold in general when G is non-compact and V is infinite-dimensional. The main issue is that there may then exist nontrivial representations having no irreducible subrepresentations. For instance, the (right) regular representation of $\mathbb{R}$ on $L^2(\mathbb{R})$ has the property. A basic result of Gelfand–Graev–Piatetski-Shapiro shows that a sufficient condition on the representation for it to decompose as a sum of irreducibles is that it be unitary and of compact type (Definition 64). This is very handy in applications involving non-compact groups. Recall that "subrepresentation" means "closed invariant subspace."

Theorem 70. *Let G be locally compact and unimodular, and let (π, V) be a unitary Hilbert representation of compact type. Then V decomposes as a Hilbert direct sum of irreducible subrepresentations V_j, each occurring with finite multiplicity (i.e., for each j, the number of k with $V_j \cong V_k$ is finite).*

Proof. By Zorn's lemma, we may find a maximal collection (V_j) of mutually orthogonal irreducible subrepresentations V_j of V. We must show that $\oplus V_j$ is dense in V. If not, then its orthogonal complement V' is a nonzero representation, satisfying the same hypotheses as V, which contains no irreducible subrepresentations. Replacing V by V' if necessary, our task is thus to show that if V is nonzero, then it contains at least one irreducible subrepresentation.

The results of §5.4 (applied to any nonzero vector $v \in V$, and with ε taken sufficiently small in terms of v) imply that there exists $f \in C_c(G)$, real-valued and with $f(g) = f(g^{-1})$, so that the integral operator $\pi(f)$ is not identically zero. Since π is assumed of compact type, the operator $\pi(f)$ is compact. By construction, it is self-adjoint. By the spectral theory of self-adjoint compact operators, we may find a nonzero real number λ so that the λ-eigenspace V_λ of $\pi(f)$ is nonzero and finite-dimensional.

Take any nonzero $v \in V_\lambda$, let Gv denote its orbit under G, and let $\langle Gv \rangle$ denote the closure of the span of Gv. Then $\langle Gv \rangle$ is a subrepresentation of V; it is the smallest subrepresentation containing v.

---

[9]End of half-lecture #8, Thursday, 14 March

The key observation is that any decomposition of the representation $\langle Gv \rangle$ descends to a decomposition of the vector v inside the finite-dimensional eigenspace V_λ. Indeed, if $\langle Gv \rangle$ is not reducible, then it contains a nonzero proper subrepresentation W_1, whose orthogonal complement W_1' in $\langle Gv \rangle$ is likewise invariant, giving a decomposition $\langle Gv \rangle = W_1 \oplus W_1'$. The subspaces in this decomposition are invariant by G, hence also by $T := \pi(f) - \lambda$, so writing $v = v_1 + v_1'$ with $v_1 \in W_1$ and $v_1' \in W_1'$, we deduce from the identity $Tv = 0$ that in fact $Tv_1 = Tv_1' = 0$, hence that $v_1, v_1' \in V_\lambda$.

We know that v_1 and v_1' are both nonzero, because if (say) $v_1' = 0$, then v belongs to the proper subrepresentation W_1 of $\langle Gv \rangle$, contrary to the construction of the latter. Since v_1' belongs to V_λ and to $\langle Gv \rangle$ but not to W_1, we have the strict containment

$$V_\lambda \cap \langle Gv_1 \rangle \subseteq V_\lambda \cap W_1 \subsetneq V_\lambda \cap \langle Gv \rangle \tag{37}$$

between finite-dimensional spaces. We now repeat the same construction but with v replaced by v_1. If $\langle Gv_1 \rangle$ is reducible, then we may decompose it as $W_2 \oplus W_2'$ and likewise v_1 as $v_2 + v_2'$, with $v_2, v_2' \in V_\lambda$. If $\langle Gv_2 \rangle$ is reducible, then we may decompose $v_2 = v_3 + v_3'$. Thanks to (37), this procedure gives us an irreducible subrepresentation after at most $\dim(V_\lambda \cap \langle Gv \rangle)$ iterations.

The slicker way to write the proof is of course to assume from the outset that v was chosen to minimize the dimension of $V_\lambda \cap \langle Gv \rangle$; reducibility of $\langle Gv \rangle$ then gives a contradiction.

For the finiteness of multiplicity, we can find for each V_j a self-adjoint integral operator $\pi(f)$ having some nonzero eigenvalue λ in V_j. The same eigenvalue then shows up in every V_k that is isomorphic to V_j. Since $\pi(f)$ is assumed compact, its eigenspaces are finite-dimensional, and so the number of such V_k is finite. $\qquad\square$

6. Algebraicity of compact Lie groups

Our goal is to explain the meaning and proof of the phrase "every compact Lie group is algebraic." This serves both as an application of the theory developed so far and as a tool for further study also of non-compact groups.

6.1. Preliminaries on algebraic groups.
First, we need to "review" some basics on algebraic groups.

Definition 71. Let k be an infinite field and $n \geq 0$. For a collection $S \subseteq k[x_1, \ldots, x_n]$ of polynomials, the *vanishing locus* $V(S)$ is the set

$$V(S) := \{p \in k^n : f(p) = 0 \text{ for all } f \in S\}$$

of common zeros. An *algebraic subset* of k^n is a subset of the form $V(S)$ for some S; that is to say, it is a subset defined by polynomial equations.

The *vanishing ideal* of a subset X of k^n is given by

$$I(X) := \{f \in k[x_1, \ldots, x_n] : f(p) = 0 \text{ for all } p \in X\}.$$

The *coordinate ring* of an algebraic subset $X \subseteq k^n$ is the ring

$$k[X] := k[x_1, \ldots, x_n]/I(X),$$

which may be regarded as the space of "polynomial functions from X to k." The *Zariski topology* on k^n (or any algebraic subset thereof) is that for which the closed subsets are the algebraic sets.

We note that Hilbert's basis theorem implies that for each S there exists a finite subset $S_0 \subset S$ so that $V(S) = V(S_0)$, so that any algebraic set may be defined by finitely-many polynomial equations.

It's not hard to check that the "Zariski topology" as described above defines an actual topology. By comparison, a basic theorem in topology says that in a nice enough topological space X (e.g., a compact metric space), for any closed subset $Z \subseteq X$ there exists a continuous function $f : X \to \mathbb{R}$ so that $Z = \{p \in X : f(p) = 0\}$, hence that the closed subsets of X are precisely those that can be defined by *continuous* equations. The analogy with the Zariski topology should be clear.

Morphisms between algebraic sets $X \subseteq k^m$ and $Y \subseteq k^n$ are maps $X \to Y$ defined by polynomials in the coordinates. They are obviously Zariski continuous. We obtain in this way a category, with objects the algebraic sets and morphisms the morphisms. An *affine variety* is basically the same thing as an algebraic set, but (by some conventions) without emphasis on any particular embedding in some k^n.

Given two sets $X \subseteq k^m$ and $Y \subseteq k^n$, we may define their product $X \times Y \subseteq k^{m+n}$. If X and Y are algebraic, then so is $X \times Y$: if $X = V(S_1)$ and $Y = V(S_2)$, then $X \times Y = V(\{f_1 \otimes f_2 : f_1 \in S_1, f_2 \in S_2\})$, where $(f_1 \otimes f_2)(p, q) := f_1(p)f_2(q)$ for $p \in k^m, q \in k^n$. We may thus speak of the Zariski topology on $X \times Y$ with respect to k^{m+n}. We caution that this is not in general the same as the product of the Zariski topologies on X and Y.

Given an algebraic set X, we'll sometimes abuse notation slightly by defining $V(S)$ for a subset S of $k[X]$ to be the set of common zeros inside X of elements of S, and, for a subset Y of X, by $I(Y)$ the ideal in $k[X]$ consisting of f that vanish on Y. Any such abuse should be clear by context.

As a basic example of reasoning with these definitions, we verify the following:

Lemma 72. *For any subset $X \subseteq k^n$, the Zariski closure $\mathrm{Zcl}(X)$ is equal to $V(I(X))$.*

Proof. Well, $V(I(X))$ contains X ("X vanishes under any polynomial that vanishes on X"), and $V(I(X))$ is algebraic, hence Zariski closed. Conversely, let Y be any Zariski closed set containing X, thus $Y = V(T)$ for some $T \subseteq k[x_1, \ldots, x_n]$. Since $X \subseteq V(T)$ ("X vanishes under every element of T"), we have $T \subseteq I(X)$ ("every element of T vanishes on X"), and so $V(T) \supseteq V(I(X))$ ("the more equations, the fewer solutions"). Thus $V(I(X))$ is the smallest Zariski closed set containing X, as required. $\qquad\square$

The group $\mathrm{GL}_n(k)$ of invertible matrices is not obviously algebraic, because it is defined inside the ambient space of matrices $M_n(k) \cong k^{n^2}$ by the polynomial *inequation* $\det \neq 0$ rather than by a system of polynomial *equations*. We "make it algebraic" by using the embedding

$$\mathrm{GL}_n(k) \cong \{(x, y) \in M_n(k)^2 : xy = 1\} \subseteq M_n(k)^2 \cong k^{2n^2}$$

$$g \mapsto (g, g^{-1}),$$

to view $\mathrm{GL}_n(k)$ as an algebraic subset of k^{2n^2}. Then the coordinate ring $k[\mathrm{GL}_n(k)]$ consists of functions $\mathrm{GL}_n \to k$ given by polynomials in the entries of a matrix together with its inverse.

Definition 73. An *algebraic group* is an abstract subgroup of $\mathrm{GL}_n(k)$ that is also an algebraic set.

When $k = \mathbb{C}$, we speak of a complex algebraic group. When $k = \mathbb{R}$, we speak of a real algebraic group. Note that we may think of $\mathrm{GL}_n(\mathbb{C})$ as a real algebraic subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$. Note also that algebraic groups over $k = \mathbb{R}$ or $\mathbb{C}$ are Lie groups (e.g., by the general theorem that closed subgroups of matrix groups are Lie groups).

Definition 73 is a bit *ad hoc*. In much the same way that we defined compact Lie groups first without reference to any matrix embedding and then showed as a consequence (Corollary 58) of the Peter–Weyl theorem that such an embedding exists, one might more properly define an algebraic group to be an algebraic set equipped with a group law for which multiplication and inversion are described in coordinates via polynomials. Arguments similar to those in the proof of Corollary 58 then confirm that this apparently more general definition is ultimately equivalent to what we've given here.

Lemma 74. *If G is an algebraic group and $H < G$ is an abstract subgroup, then the Zariski closure $\mathrm{Zcl}(H)$ of H inside G is an algebraic subgroup of G.*

Proof. We just need to check that $\mathrm{Zcl}(H)$ is closed under multiplication and inversion. We check this for multiplication, leaving the case of inversion to the reader (alternatively, one could get fancy and consider maps like $(p, q) \mapsto pq^{-1}$). Let $p, q \in \mathrm{Zcl}(H)$, and let $pq \in U \subseteq G$ be a neighborhood. Since matrix multiplication is described by polynomials, it is Zariski continuous, and so we can find a neighborhood $p \in V_1 \subseteq G$ so that $p'q \in U$ for all $p' \in V_1$. Since p belongs to the closure of H, we know that V_1 intersects H, so choose some $p' \in V_1 \cap H$. Then $p'q \in U$, so by the same argument, we can find a neighborhood $q \in V_2 \subseteq G$ so that $p'q' \in U$ for all $q' \in V_2$. Since q is in the closure of H, we know that V_2 intersects H, so we may find some $q' \in V_2 \cap H$. Since H is a subgroup, we then have $p'q' \in U \cap H$, so U intersects H. Since U was arbitrary, we conclude that $pq \in \mathrm{Zcl}(H)$. $\qquad\square$

6.2. **Complexification of a compact Lie group.** Now let K be a compact subgroup of $\mathrm{GL}_n(\mathbb{C})$, hence a Lie subgroup. Recall also, by Corollary 58, that every compact Lie group is of this form. Since the standard representation $K \hookrightarrow \mathrm{GL}_n(\mathbb{C})$ is unitarizable, we know that K is conjugate to a subgroup of $\mathrm{U}(n)$. Let's assume for convenience that in fact $K \leq \mathrm{U}(n)$. Let G denote the complex Zariski closure of K in $\mathrm{GL}_n(\mathbb{C})$, thus

$$G = \{g \in \mathrm{GL}_n(\mathbb{C}) : f(g) = 0 \text{ for all } f \in \mathbb{C}[\mathrm{GL}_n(\mathbb{C})] \text{ with } f|_K = 0\},$$

where as before $\mathbb{C}[\mathrm{GL}_n(\mathbb{C})]$ denotes the space of functions $f : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}$ given by polynomials in the entries of a matrix together with its inverse. (For example, if $K = \mathrm{U}(n)$, then $G = \mathrm{GL}_n(\mathbb{C})$.) We denote by

$$\Theta : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$$

the "inverse conjugate transpose" map

$$\Theta(g) := {}^t\overline{g}^{-1} = (g^*)^{-1}, \quad g^* := {}^t\overline{g}$$

and by

$$\theta : \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$$
$$\theta(x) := -{}^t\overline{x} = -x^*, \quad x^* := {}^t\overline{x}$$

its differential. Observe that $\mathrm{U}(n) = \{g \in \mathrm{GL}_n(\mathbb{C}) : \Theta(g) = g\}$ and $\mathfrak{u}(n) = \{x \in \mathfrak{gl}_n(\mathbb{C}) : \theta(x) = x\}$.

Theorem 75. *With notation and assumptions as above:*
  (i) *G is a complex algebraic group that is closed under Θ.*
  (ii) *The map $\mathbb{C}[G] \to \{$functions $K \to \mathbb{C}\}$ given by restriction $f \mapsto f|_K$ induces an isomorphism $\mathbb{C}[G] \cong \mathcal{A}(K)$.*
 (iii) *$K = G \cap \mathrm{U}(n)$. In particular, K is a real algebraic group.*[10]
  (iv) *K is a maximal compact subgroup of G.*
  (v) *$\mathfrak{g}$ is closed under θ. We have $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{u}(n) = \{x \in \mathfrak{g} : \theta(x) = x\}$. Setting $\mathfrak{p} := \{x \in \mathfrak{g} : \theta(x) = -x\}$, we have $\mathfrak{p} = i\mathfrak{k}$ and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$.*

Proof.   (i) By lemma 74, G is a complex algebraic group. Since Θ fixes $\mathrm{U}(n)$ pointwise and $\mathrm{U}(n)$ contains K, we have in particular $\Theta(K) = K$. It follows that the space $I(K)$ of polynomials $f \in \mathbb{C}[\mathrm{GL}_n(\mathbb{C})]$ that vanish on K is closed under the map

$$f \mapsto \overline{f \circ \Theta} : g \mapsto \overline{f(\Theta(g))} = \overline{f({}^t\overline{g}^{-1})},$$

hence that $G = V(I(K))$ is closed under Θ.

  (ii) We observe first (as in the discussion of §4.9) that $f|_K$ belongs to $\mathcal{A}(K)$: indeed, $\mathbb{C}[G]$ is the union over $d \geq 0$ of the subspaces defined by polynomials of degree bounded by d, each of which is finite-dimensional and K-invariant. We thus get a well-defined map of $\mathbb{C}$-algebras $\mathbb{C}[G] \to \mathcal{A}(K)$. The injectivity of this map is clear from the definitions of G and of $\mathbb{C}[G] = \mathbb{C}[\mathrm{GL}_n(\mathbb{C})]/I(K)$: if $f|_K = 0$, then $f \in I(K)$, i.e., $f = 0$ as an element of $\mathbb{C}[G]$. We henceforth identify $\mathbb{C}[G]$ with its image in $\mathcal{A}(K)$. It remains to show that in fact $\mathbb{C}[G] = \mathcal{A}(K)$. To that end, observe first that $\mathbb{C}[G]$ is invariant under the action of $K \times K$ on $\mathcal{A}(K)$ by left and right translation (using here that matrix multiplication is described by polynomials and that K is a subgroup of G). Recall next that $\mathcal{A}(K)$ is an orthogonal direct sum of finite-dimensional subspaces $\mathcal{A}(\pi)$, taken

---
[10]End of half-lecture #9, Tuesday, 19 March

over $\pi \in \mathrm{Irr}(K)$. Note finally, thanks to the identity $g = \Theta(g)$ satisfied by every $g \in K$, that $\mathbb{C}[G]$ contains all polynomials in the real and imaginary parts of the entries of a matrix $g \in K$ together with its inverse. It follows by Stone-Weierstrass (as in the discussion of §4.9) that $\mathbb{C}[G]$ is dense in $\mathcal{A}(K)$ with respect to the topology of $C(G)$.

To complete the proof, we just need to check that any $K \times K$-invariant dense subspace of $\mathcal{A}(K)$ is actually equal to $\mathcal{A}(K)$. Maybe there's a simpler way to see this, but let's see. We claim that for any finite subset Π of $\mathrm{Irr}(K)$ and any $K \times K$-invariant subspace V of $\oplus_{\pi \in \Pi} \mathcal{A}(\pi)$, one has $V = \oplus_{\pi \in \Pi_0} \mathcal{A}(\pi)$ for some subset Π_0 of Π. (To see that the claim suffices, note that for each $\pi \in \mathrm{Irr}(K)$, we may find $f \in \mathbb{C}[G]$ so that the component $f_\pi \in \mathcal{A}(\pi)$ of f is nonzero, because otherwise $\mathbb{C}[G]$ would be contained in the proper closed subspace of $\mathcal{A}(K)$ given by the orthogonal complement of $\mathcal{A}(\pi)$; then take for Π the set of all π' for which $f_{\pi'} \neq 0$ and for V the span of the $K \times K$-orbit of f. It follows from the claim that then $\mathbb{C}[G]$ contains $\mathcal{A}(\pi)$. Since π was arbitrary, we conclude that $\mathbb{C}[G] = \mathcal{A}(K)$.) The claim follows immediately from character theory for $K \times K$, which gives that $\chi_V = \sum_{\pi \in \Pi} n(\pi) \overline{\chi_\pi} \otimes \chi_\pi$ for some $n(\pi) \in \{0,1\}$; here $(\overline{\chi_\pi} \otimes \chi_\pi)(g_1, g_2) := \overline{\chi_\pi}(g_1) \chi_\pi(g_2)$.

(iii) Set $M := G \cap \mathrm{U}(n)$, so that $K \leq M \leq G$. Then M is compact, and G is also the Zariski closure of M, so by what we've already shown, the map $\mathbb{C}[G] \to \mathcal{A}(M)$ is an isomorphism, hence in particular has dense image in $C(M)$. It follows readily that $M = K$. Indeed, suppose otherwise that M is strictly larger than K. We let M act on $C(M)$ and $\mathcal{A}(M)$ by the right regular representation, and use superscripts as in $C(M)^K$ to denote the fixed subspace, as usual. We argue as follows:

- $C(M)^K$ *contains non-constant functions.* Indeed, take any nonzero non-negative $f \in C(M)$ that vanishes on K, and consider the function $M \ni x \mapsto \int_{k \in K} f(xk)$, where here and henceforth $\int$ denotes an integral with respect to the probability Haar.

- $\mathbb{C}[G]^K$ *contains non-constant functions.* We have seen that $\mathbb{C}[G]$ maps isomorphically to $\mathcal{A}(M)$, hence has dense image in $C(M)$. Take $f \in C(M)^K$ non-constant, thus $f(x_1) \neq f(x_2)$ for some $x_1, x_2 \in M$. Choose $\varepsilon > 0$ so that $|f(x_1) - f(x_2)| \geq 3\varepsilon$. Choose $f' \in \mathbb{C}[G]$ so that $\|f - f'\|_\infty \leq \varepsilon$. Set $f''(x) := \int_{k \in K} f'(xk)$. Then $f'' \in \mathbb{C}[G]^K$. By the right K-invariance of f, we have $(f - f'')(x) = \int_{k \in K} (f - f')(xk)$, hence by the triangle inequality we have $\|f - f''\|_\infty \leq \|f - f'\|_\infty \leq \varepsilon$. By another application of the triangle inequality, we deduce that $|f''(x_1) - f''(x_2)| \geq \varepsilon$. Thus f'' is non-constant.

- *We obtain a contradiction.* We have seen that $\mathbb{C}[G]$ injects into $\mathcal{A}(K)$, hence likewise $\mathbb{C}[G]^K$ into $\mathcal{A}(K)^K$, but every element of $\mathcal{A}(K)^K$ is manifestly constant.

(iv) If M is a larger compact subgroup, then we may assume (after conjugating) that it is contained in $\mathrm{U}(n)$. Then both M and K are contained in $\mathrm{U}(n)$ and have G as their Zariski closure, so by (iii), we have $K = G \cap \mathrm{U}(n) = M$.

(v) The identity $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{u}(n)$ follows from part (iii). Since $\theta^2 = 1$, we can decompose $\mathfrak{g}$ into its ± 1 eigenspaces for θ, i.e., $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$; explicitly, $x = x^+ + x^-$, where $x^+ \in \mathfrak{k}, x^- \in \mathfrak{p}$ are given by $x^\pm := (x \pm \theta(x))/2$. Since θ is anti-linear, multiplication by i induces an isomorphism $x^+ \cong x^-$, i.e., $\mathfrak{p} = i\mathfrak{k}$.

$$\square$$

Remark 76. It wasn't essential to refer to the group $\mathrm{U}(n)$ here. The definition of G doesn't involve it. The involution $f \mapsto \overline{f \circ \Theta}$ on $\mathbb{C}[G]$ corresponds under the isomorphism $\mathbb{C}[G] \cong \mathcal{A}(K)$ to the involution σ on $\mathcal{A}(K)$ given by $\sigma(f)(k) = \overline{f(k)}$ for all $k \in K$.

6.3. **A glimpse of Tannakian duality.** Here we address the question: how can we describe a compact Lie group K in terms of its coefficient ring $\mathcal{A}(K)$? We warm-up with some motivating analogues.

- For an algebraic set $X \subseteq k^n$ (over an infinite field k) with coordinate ring $k[X] = k[x_1, \ldots, x_n]/I(X)$, we have a bijection

$$X \leftrightarrow \mathrm{Hom}_k(k[X], k)$$

$$p \mapsto [f \mapsto f(p)]$$

$$(\ell(x_1), \ldots, \ell(x_n)) \leftarrow\!\shortmid \ell$$

  between points of X and functionals on the coordinate ring. Moreover, we can reconstruct X from the abstract k-algebra $k[X]$ by choosing a system of generators $x_1, \ldots, x_n$.
- For an algebraic group G over k, we can recover G as an algebraic set from $\mathrm{Hom}_k(k[G], k)$, as above, but we'd also like to keep track of the group law $G \times G \to G$. It corresponds to the pullback map

$$\Delta : k[G] \to k[G \times G] \cong k[G] \otimes k[G],$$

  called *comultiplication*, and characterized by the identity

$$\Delta(f)(g_1, g_2) = f(g_1 g_2) \text{ for all } f \in k[G] \text{ and } g_1, g_2 \in G.$$

Now let $K \leq \mathrm{U}(n)$ be a compact Lie group, with Zariski closure $G \leq \mathrm{GL}_n(\mathbb{C})$ as above. Then

$$G \cong \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}[G], \mathbb{C}) \cong \mathrm{Hom}_{\mathbb{C}}(\mathcal{A}(K), \mathbb{C}). \tag{38}$$

The comultiplication Δ on $\mathbb{C}[G]$ corresponds under the isomorphism with $\mathcal{A}(K)$ to the map

$$\Delta : \mathcal{A}(K) \to \mathcal{A}(K) \otimes \mathcal{A}(K)$$

described explicitly in terms of matrix coefficients by the usual matrix multiplication rule

$$\Delta(\pi_{ij}) = \sum_k \pi_{ik} \otimes \pi_{kj}$$

and in basis-free manner as the map induced by the maps $\mathrm{End}(V)^* \to \mathrm{End}(V)^* \otimes \mathrm{End}(V)^*$ (V a finite-dimensional representation of G) coming from the multiplication maps $\mathrm{End}(V) \otimes \mathrm{End}(V) \to \mathrm{End}(V)$.

These observations already show that G is determined as an algebraic group by the pair $(\mathcal{A}(K), \Delta)$; in particular, G is independent of the choice of linear embedding used in its construction. We will henceforth refer to G as the *complexification* of K.

We can say something similar about K:

Theorem 77. *Let K be a compact Lie group. Then the map $k \mapsto [f \mapsto f(k)]$ defines a bijection*

$$K \leftrightarrow \{\ell \in \mathrm{Hom}_{\mathbb{C}}(\mathcal{A}(K), \mathbb{C}) : \ell(\sigma(f)) = \overline{f} \text{ for all } f \in \mathcal{A}(K)\}, \qquad (39)$$

where $\sigma : \mathcal{A}(K) \to \mathcal{A}(K)$ is given by complex conjugation. The group law on K is described by Δ, i.e., for all $f \in \mathcal{A}(K)$ and $k_1, k_2 \in K$, we have $f(k_1 k_2) = \Delta(f)(k_1, k_2)$.

Informally, "$(\mathcal{A}(K), \Delta)$ determines K."

Proof. Let G be as constructed above, and let ℓ be an element of the RHS of (39). By (38), we may write $\ell(f) = f(g)$ for some $g \in G$. By Remark 76, we have $\overline{f(\Theta(g))} = f(g)$ for all $f \in \mathbb{C}[G]$, hence $\Theta(g) = g$. By Theorem 75, we conclude that $g \in K$. $\qquad \square$

There's much more to say about this topic than we will here. Some keywords: Tannaka–Krein duality, Hopf algebras.

6.4. **Cartan decomposition.** Set

$$P(n) := \{g \in \mathrm{GL}_n(\mathbb{C}) : p^* = p, \text{ positive-definite}\},$$

$$\mathfrak{p}(n) := \{x \in \mathfrak{gl}_n(\mathbb{C}) : x^* = x\}.$$

Then the map $\exp : \mathfrak{p}(n) \to P(n)$ is a bijection; indeed, by the spectral theorem for hermitian matrices, we have for each $x \in \mathfrak{p}(n)$ and $p \in P(n)$ that

$$x \sim \begin{pmatrix} x_1 & & \\ & \cdots & \\ & & x_n \end{pmatrix}, \quad p \sim \begin{pmatrix} p_1 & & \\ & \cdots & \\ & & p_n \end{pmatrix}$$

for some $x_i \in \mathbb{R}, p_j \in \mathbb{R}_{>0}$, where $\sim$ denotes conjugacy. We deduce readily that $\exp : \mathfrak{p}(n) \to P(n)$ and $\log : P(n) \to \mathfrak{p}(n)$ (with the latter defined by the "functional calculus," i.e., by taking the logarithm of each diagonal entry) define mutually inverse bijections. Note that for $p \in P(n)$, we can define p^t for any complex number t, either via the logarithm as $\exp(t \log p)$ or directly via the functional calculus.

The following might be known from linear algebra:

Lemma 78 (Polar decomposition). *The map $\mathrm{U}(n) \times P(n) \to \mathrm{GL}_n(\mathbb{C})$ given by $(k, p) \mapsto kp$ is bijective.*

Proof. Injectivity: if $g = kp$, then $g^* g = p^* k^* k p = p^2$, so $p = \sqrt{p^2} := \exp(\frac{1}{2} \log p)$ and thus also $k = gp^{-1}$ are determined by g. Surjectivity: observe that $g^* g \in P(n)$ and define $p := \sqrt{g^* g} \in P(n)$, $k := gp^{-1} \in \mathrm{U}(n)$. $\qquad \square$

Lemma 79. The map

$$\mathrm{U}(n) \times \mathfrak{p}(n) \to \mathrm{GL}_n(\mathbb{C})$$

$$(k, x) \mapsto k \exp(x)$$

is a diffeomorphism.

Proof. We have already checked the bijectivity. The forward map is smooth. We need to know that the inverse is smooth. This boils down to the fact that $\exp : \mathfrak{p}(n) \to P(n)$ is submersive, which is a consequence of the following lemma and the fact that for $x \in \mathfrak{p}(n)$, the eigenvalues of ad_x are real. $\qquad \square$

The following is part of the BCHD law, and describes the derivative of the exponential map on any Lie group.

Lemma 80. *Let G be any Lie group, Fix $x \in \mathfrak{g}$. Let $z \in \mathfrak{g}$ be small enough. Let $y \in \mathfrak{g}$ be the small element for which*

$$\exp(x + z) = \exp(x) \exp(y).$$

Then

$$y = F(\mathrm{ad}_x)z + \mathrm{O}(|z|^2),$$

where

$$F(t) := \frac{1 - e^{-t}}{t} = 1 - \frac{t}{2!} + \frac{t^2}{3!} - \cdots .$$

In particular, $\exp$ *is a local diffeomorphism at x iff $F(\mathrm{ad}_x)$ is invertible iff ad_x has no eigenvalues of the form $2\pi i k$, with $0 \neq k \in \mathbb{Z}$.*

Proof. (Omitted from lecture.) For $t \in [0, 1]$, set $f(t) := \log(\exp(-tx) \exp(t(x + z)))$, so that $f(0) = 0$ and $f(1) = y$. We compute that for small $\varepsilon > 0$,

$$f(t + \varepsilon) = \mathrm{Ad}_{-tx}(\varepsilon z) + tz + \mathrm{O}(\varepsilon |z|^2 + \varepsilon^2),$$

where $\mathrm{Ad}_x := e^{\mathrm{ad}_x} = \mathrm{Ad}(e^x)$. (To see this, first write $\exp(-(t + \varepsilon)x) \exp((t + \varepsilon)(x + z)) = \exp(-tx) \exp(-\varepsilon x) \exp(\varepsilon(x + z)) \exp(t(x + z))$, then approximate $\exp(-\varepsilon x) \exp(\varepsilon(x+z)) = \exp(\varepsilon z + \mathrm{O}(\varepsilon^2))$ and deduce that $f(t+\varepsilon) = \log(\exp(\mathrm{Ad}_{-tx}(\varepsilon z) + \mathrm{O}(\varepsilon^2)) \exp(\mathrm{O}(z))) = \mathrm{Ad}_{-tx}(\varepsilon z) + \mathrm{O}(\varepsilon |z|^2 + \varepsilon^2)$, as required.) Thus

$$f'(t) = \mathrm{Ad}_{-tx}(z) + \mathrm{O}(|z|^2)$$

and so

$$y = \int_{t=0}^{1} e^{-t \, \mathrm{ad}_x} z \, dt + \mathrm{O}(|z|^2),$$

which leads to the required conclusion by the formula $\int_{t=0}^{1} e^{-tu} \, dt = F(u)$.   □

Theorem 81. *Let K be a compact Lie group, with complexification G. Write $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ as in Theorem 75. Then the map*

$$K \times \mathfrak{p} \to G$$

$$(k, x) \mapsto k \exp(x)$$

is a diffeomorphism. In particular,

- K *is a deformation retract of G,*
- K *meets every connected component of G,*
- K *is connected if and only if G is connected,*

and so on.

Proof. In view of Lemma 79, we just need to check that if $g \in G$ is written $g = kp$ with $k \in \mathrm{U}(n)$ and $p \in P(n)$, then in fact $p \in P := \exp(\mathfrak{p})$; it follows then $p \in G$, hence that $k = gp^{-1} \in G \cap \mathrm{U}(n) = K$. It will suffice to show that $p^t \in G$ for all $t \in \mathbb{R}$, because then $\partial_{t=0} p^t$ belongs to $\mathfrak{g} \cap \mathfrak{p}(n) = \mathfrak{p}$. Since $G = V(I(G))$, it will suffice to show for each $f \in I(G)$ that $f(p^t) = 0$ for all $t \in \mathbb{R}$. Since $p^2 = g^* g$ and G is a group, we know that $f(p^t) = 0$ for all $t \in 2\mathbb{Z}$. Let $e^{x_1}, \ldots, e^{x_n}$ denote the diagonal entries of p with respect to some basis. Then $f(p^t)$ is a polynomial in the quantities $e^{\pm t x_1}, \ldots, e^{\pm t x_n}$. Collecting common exponents, we may write

$$f(p^t) = \sum_{\gamma \in \Gamma} c_\gamma e^{\gamma t}$$

for some finite subset Γ of $\mathbb{R}$ and some nonzero complex coefficients c_γ. Suppose for the sake of contradiction that $f(p^t)$ is nonzero. Then Γ is nonempty. Let $\gamma \in \Gamma$ be the largest element. Then $\lim_{t \to \infty} e^{-\gamma t} f(p^t) = c_\gamma \neq 0$. But $f(p^t) = 0$ for all $t \in 2\mathbb{Z}$, giving the required contradiction. $\qquad \square$

6.5. Algebraic representations.

Definition 82. A *morphism* of algebraic groups is a group homomorphism that is also a morphism of the underlying algebraic sets, i.e., a map described in coordinates by polynomials.

Let G be an algebraic group over an infinite field k. An *algebraic representation* of G is a finite-dimensional k-vector space V equipped with a morphism of algebraic groups $\pi : G \to \mathrm{GL}(V)$.

Theorem 83. *Let K be a compact Lie group. Let G denote its complexification. Then algebraic representations (σ, W) of G are in natural bijection with finite-dimensional representations (π, V) of K. The bijection sends V to W and W to V. To get from σ to π, one restricts. To get from π to σ, one writes π in matrix form as (π_{ij}), with $\pi_{ij} \in \mathcal{A}(K)$; one then uses the isomorphism $\mathcal{A}(K) \cong \mathbb{C}[G]$ to identify each π_{ij} with a regular function on G, and takes $\sigma = (\pi_{ij})$ the algebraic representation whose matrix entries are given by those regular functions.*

Proof. The reader is encouraged to check that this follows readily from Theorem 75. $\qquad \square$

[11]

6.6. Reductive groups.

Which groups G arise as in Theorem 75, i.e., as the Zariski closure of a compact subgroup of $\mathrm{GL}_n(\mathbb{C})$? Recall that an element $g \in \mathrm{GL}_n(\mathbb{C})$ is *unipotent* if $(g-1)^n = 0$. We say that a subgroup of $\mathrm{GL}_n(\mathbb{C})$ is *unipotent* if each of its elements is unipotent.

Definition 84. A complex algebraic group $G \leq \mathrm{GL}_n(\mathbb{C})$ is called *reductive* if it contains no nontrivial (i.e., other than $\{1\}$) normal unipotent subgroups.

For instance, the group $\begin{pmatrix} 1 & \mathbb{C} \\ & 1 \end{pmatrix}$ is not reductive, because the whole group is unipotent.

Theorem 85. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a complex algebraic group. The following are equivalent:*

   (i) G is reductive.
   (ii) G has a Zariski dense compact subgroup K.
   (iii) G is conjugate to a group that is closed under $\Theta : g \mapsto {}^t\overline{g}^{-1}$.
   (iv) Every algebraic representation V of G is completely reducible, i.e., decomposes as a direct sum $\oplus W_i$ of invariant irreducible subspaces.

Proof. The proof that (i) implies (iii) implies (ii) is lengthy, and will not be discussed in detail. The basic idea is to complete most of the classifications of reductive algebraic groups and of compact Lie groups, and to observe that the same objects (root data) arise in both classifications. We will soon discuss these ideas for compact Lie groups; a parallel discussion applies in the algebraic setting.

---

[11]End of lecture #10, Tuesday, 26 March

That (ii) implies (iii) follows from Theorem 75.

Let's assume (ii) and deduce (iv). We have seen (Theorem 22) that V decomposes as a direct sum $\oplus W_i$ of K-invariant K-irreducible subspaces. We claim that if a subspace $W \subseteq V$ of V is K-invariant, then it is likewise G-invariant. It's obvious that any K-irreducible G-invariant subspace is likewise G-irreducible, so the claim suffices. To verify the claim, the point is that leaving a subspace invariant is an algebraic condition, which thus extends Zariski-continuously from K to G. In more detail, consider the set $X := \{g \in \mathrm{GL}(V) : gW \subseteq W\}$. Then X is algebraic; for instance, if we extend a basis $e_1, \ldots, e_m$ for W to a basis $e_1, \ldots, e_n$ for V and denote by $e_1^*, \ldots, e_n^*$ the corresponding dual basis, then $X = \{g : e_j^*(ge_i) = 0$ for $i = 1..m, j = m+1..n\}$. Let $\pi : G \to \mathrm{GL}(V)$ denote the action map. Then $\pi^{-1}(X)$ is algebraic. Since W is K-invariant, we have $K \subseteq \pi^{-1}(X)$. Since K is Zariski dense in G, it follows that $G \subseteq \pi^{-1}(X)$, hence that X is G-invariant, as required.

Let's finally assume (iv) and deduce (i). Let N be a normal unipotent subgroup. We must show that $N = \{1\}$. Let $V = \mathbb{C}^n$ denote the standard representation of G, thus $G \subseteq \mathrm{GL}(V)$. By assumption, $V = \oplus W_i$ with each W_i an irreducible subrepresentation. By Engel's theorem (or perhaps a variant), the fixed space $W_i^N = \{w \in W_i : nw_i = w_i$ for all $n \in N\}$ is nontrivial. Since N is normal, W_i^N is G-invariant. (Indeed, if $v \in W_i^N$, $g \in G$ and $n \in N$, then $n' := g^{-1}ng \in N$, so $ngv = gn'v = gv$, hence $gv \in W_i^N$.) Since W_i is irreducible, it follows that $W_i = W_i^N$, hence that $V = V^N$. This says that every element n of N fixes every element of V, i.e., that $n = 1$, i.e., that $N = \{1\}$, as required. $\qquad\square$

6.7. **Unitary trick.** We record a further variant of the above considerations. We focus on an example; what's relevant here is that $\mathrm{SL}_n(\mathbb{C})$ and $\mathrm{SU}(n)$ are connected and simply-connected. Note that by the Cartan decomposition (§6.4), knowing either of these properties for one of these groups implies the same property for the other group.

Theorem 86. *Let V be a finite-dimensional complex vector space. The following are all in natural bijection:*

(1) Algebraic representations $\mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}(V)$
(2) Holomorphic representations $\mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}(V)$ (i.e., morphisms of complex Lie groups)
(3) Holomorphic representations $\mathfrak{sl}_n(\mathbb{C}) \to \mathfrak{gl}(V)$ (i.e., morphisms of complex Lie algebras)
(4) Representations $\mathrm{SU}(n) \to \mathrm{GL}(V)$
(5) Representations $\mathfrak{su}(n) \to \mathrm{GL}(V)$
(6) Representations $\mathrm{SL}_n(\mathbb{R}) \to \mathrm{GL}(V)$
(7) Representations $\mathfrak{sl}_n(\mathbb{R}) \to \mathrm{GL}(V)$

These bijections are compatible with irreducibility and direct sum decompositions. Any such representation is completely reducible, i.e., decomposes as a direct sum of irreducible representations.

Proof. We go from (1) to (2) using that polynomials are holomorphic. We'll leave it as a homework problem to show that any holomorphic representation of a complex reductive group is automatically polynomial, giving the arrow from (2) to (1). (2) and (3) are equivalent by Lie theory and the fact that $\mathrm{SL}_n(\mathbb{C})$ is connected and simply-connected. (4) and (5) are likewise equivalent thanks to the analogous

properties for SU(n). (3), (5) and (7) are all equivalent thanks to the identities

$$\mathfrak{su}(n) \oplus i\,\mathfrak{su}(n) = \mathfrak{sl}_n(\mathbb{C}) = \mathfrak{sl}_n(\mathbb{R}) \oplus i\,\mathfrak{sl}_n(\mathbb{R})$$

and the fact that an $\mathbb{R}$-linear map $W \to U$ from a real vector space W to a complex vector space U extends uniquely to a $\mathbb{C}$-linear map $W_{\mathbb{C}} \to U$ from the complexification $W_{\mathbb{C}} = W \otimes_{\mathbb{R}} \mathbb{C} = W \oplus iW$. We go from (2) to (6) by restricting and from (6) to (7) by differentiating. That these bijections are compatible with irreducibility and decompositions is either clear or follows as in the proof of Theorem 85. We've seen that complete reducibility holds for representations of SU(n), so the same follows for each of the other classes of representations. $\qquad\square$

We've already (§3.7) classified the (irreducible) finite-dimensional representations of SU(n), so the stated equivalence tells us that we've implicitly classified each of the other six classes of representations.

We note that Theorems 83 and 86 are mostly independent.

We record the homework problem promised above as a theorem:

Theorem 87. *Let G be a reductive complex algebraic group and V a finite-dimensional complex vector space. Then any holomorphic representation $\pi : G \to \mathrm{GL}(V)$ is algebraic.*

Note that by contrast, non-reductive groups can have holomorphic non-algebraic representations, e.g.,

$$\begin{pmatrix} 1 & \mathbb{C} \\ & 1 \end{pmatrix} \ni \begin{pmatrix} 1 & z \\ & 1 \end{pmatrix} \mapsto \begin{pmatrix} e^z & \\ & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C}).$$

We've only scratched the tip of the iceberg here; for further reading I recommend the book by Onishchik–Vinberg.

7. STRUCTURE OF COMPACT LIE GROUPS

We aim next to show that any compact Lie group K has structure similar to that observed in §3 for the unitary groups. We'll use this to classify such K together with their representations.

7.1. **Notation related to a torus.** Let T be an n-dimensional torus. Recall from §3 that this means that T is a connected compact abelian Lie group of dimension n and that any such T is isomorphic to $U(1)^n$. We may fix such an isomorphism and regard T as the diagonal subgroup of $U(n)$. In §3, we worked with the explicit coordinates defined by this embedding. It'll be useful now to work in a more basis-free manner. This requires setting up a bit of notation.

The *character group* of T is

$$X(T) := \mathrm{Hom}(T, U(1)),$$

while the *cocharacter group* is

$$X^\vee(T) := \mathrm{Hom}(U(1), T).$$

We view these as additive groups. They are free $\mathbb{Z}$-modules of rank n. We can identify $X(T)$ and with $\mathbb{Z}^n$ by associating to $\lambda, \gamma \in \mathbb{Z}^n$ the character

$$T \ni t \mapsto t^\lambda := t_1^{\lambda_1} \cdots t_n^{\lambda_n}$$

and the cocharacter

$$U(1) \ni z \mapsto z^\gamma := \mathrm{diag}(z^{\gamma_1}, \ldots, z^{\gamma_n}).$$

There is a natural pairing

$$\langle, \rangle : X(T) \otimes_{\mathbb{Z}} X^\vee(T) \to \mathbb{Z}$$

given by observing that the composition $U(1) \to T \to U(1)$ is a morphism $U(1) \to U(1)$, and that any such morphism is of the form $z \mapsto z^k$ for some integer k (e.g., by §2.6). For $\lambda, \gamma \in \mathbb{Z}^n$, we have $(z^\gamma)^\lambda = z^{\gamma_1 \lambda_1 + \cdots + \gamma_n \lambda_n}$, so $\langle \lambda, \gamma \rangle = \gamma_1 \lambda_1 + \cdots + \gamma_n \lambda_n$.

We denote by $\mathfrak{t}$ the Lie algebra of T and by $\mathfrak{t}_{\mathbb{C}} = \mathfrak{t} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{t} \oplus i\mathfrak{t}$ its complexification. We may identify

$$\mathfrak{t}_{\mathbb{C}} = \begin{pmatrix} \mathbb{C} & & \\ & \ddots & \\ & & \mathbb{C} \end{pmatrix}.$$

We (somewhat awkwardly) set $\mathfrak{t}_{\mathbb{R}} := i\mathfrak{t}$, so that

$$\mathfrak{t}_{\mathbb{R}} = \begin{pmatrix} \mathbb{R} & & \\ & \ddots & \\ & & \mathbb{R} \end{pmatrix}. \tag{40}$$

We obtain a surjective covering map

$$\mathfrak{t}_{\mathbb{R}} \to T$$

$$x \mapsto \exp(ix).$$

We write $\mathfrak{t}_{\mathbb{C}}^* := \mathrm{Hom}_{\mathbb{C}}(\mathfrak{t}_{\mathbb{C}}, \mathbb{C}) \cong \mathbb{C}^n$ and $\mathfrak{t}_{\mathbb{R}}^* := \mathrm{Hom}_{\mathbb{R}}(\mathfrak{t}_{\mathbb{R}}, \mathbb{R}) \cong \mathbb{R}^n$. We may identify $\mathfrak{t}_{\mathbb{R}}^*$ with the subspace of $\mathfrak{t}_{\mathbb{C}}^*$ mapping $\mathfrak{t}_{\mathbb{R}}$ to $\mathbb{R}$. The identification of $\mathfrak{t}_{\mathbb{R}}^*$ with $\mathbb{R}^n$ is given by $\lambda = (\lambda_1, \ldots, \lambda_n)$ if $\lambda(x) = \lambda_1 x_1 + \cdots + \lambda_n x_n$ for $x = \mathrm{diag}(x_1, \ldots, x_n) \in \mathfrak{t}_{\mathbb{R}}$.

Each $\lambda \in X(T)$ identifies with an element $\lambda \in \mathfrak{t}_{\mathbb{R}}^*$; this identification is determined by requiring that for every $x \in \mathfrak{t}_{\mathbb{R}}$ (so that $\exp(ix) \in T$), we have

$$(\exp(ix))^\lambda = e^{i\lambda(x)}.$$

We write $\mathfrak{t}_{\mathbb{Z}}^* \subseteq \mathfrak{t}_{\mathbb{R}}^*$ for the image of $X(T)$ under this identification. The identification above of $\mathfrak{t}_{\mathbb{R}}^*$ with $\mathbb{R}^n$ then carries $\mathfrak{t}_{\mathbb{Z}}^*$ to $\mathbb{Z}^n$.

Each $\gamma \in X^\vee(T)$ identifies with an element $\gamma \in \mathfrak{t}_{\mathbb{R}}$, characterized as follows: for $\theta \in \mathbb{R}$ (so that $\exp(i\theta) \in U(1)$),

$$(\exp(i\theta))^\gamma = \exp(i\theta\gamma).$$

We denote by $\mathfrak{t}_{\mathbb{Z}} \subseteq \mathfrak{t}_{\mathbb{R}}$ the image of $X^\vee(T)$ under this identification. Then under (40), we have

$$\mathfrak{t}_{\mathbb{Z}} = \begin{pmatrix} \mathbb{Z} & & \\ & \ddots & \\ & & \mathbb{Z} \end{pmatrix}. \tag{41}$$

We note that $X(T)$ and $X^\vee(T)$ identify respectively with $\mathfrak{t}_{\mathbb{Z}}$ and $\mathfrak{t}_{\mathbb{Z}}^*$, and the pairing $X(T) \otimes X^\vee(T) \to \mathbb{Z}$ (tensor product over $\mathbb{Z}$) discussed above corresponds to the natural pairing $\langle,\rangle : \mathfrak{t}_{\mathbb{Z}} \otimes \mathfrak{t}_{\mathbb{Z}}^* \to \mathbb{Z}$ induced by the canonical duality between $\mathfrak{t}_{\mathbb{R}}$ with $\mathfrak{t}_{\mathbb{R}}^*$. In coordinates, $\langle \lambda, \gamma \rangle = \lambda_1 \gamma_1 + \cdots + \lambda_n \gamma_n$, as before.

We should note that $\mathfrak{t}_{\mathbb{Z}}$ and $\mathfrak{t}_{\mathbb{Z}}^*$ depend upon T, not just upon $\mathfrak{t}$.

Recall (§2.6) that every finite-dimensional irreducible representation of T is one-dimensional and of the form $t \mapsto t^\lambda$ for some $\lambda \in \mathfrak{t}_{\mathbb{Z}}^*$. Given a finite-dimensional representation V of T, recall (§3.3) that we may decompose

$$V = \oplus_{\lambda \in \mathfrak{t}_{\mathbb{Z}}^*} V^\lambda,$$

where $V^\lambda = \{v \in V : tv = t^\lambda v \text{ for all } t \in T\}$ is the weight space with weight λ.

7.2. Maximal tori: definition and existence.
Let K be a compact Lie group. By a *torus* $T \leq K$ we mean a closed subgroup that is a torus. Note that if T is an abstract torus and $j : T \hookrightarrow K$ is a injective immersive Lie group morphism (i.e., a "virtual Lie subgroup"), then $j(T)$ is the image under a continuous map of a compact set, hence $j(T)$ compact, hence (since K is Hausdorff) $j(T)$ is closed, and so $j(T)$ is a torus in G.

Definition 88. A *maximal torus* T of K is a torus $T \leq K$ that is not contained in any strictly larger torus in K.

Lemma 89. *Let K be a compact Lie group.*

(i) A torus $T \leq K$ is maximal if and only if $\mathfrak{t}$ is a maximal abelian subalgebra of $\mathfrak{k}$.

(ii) Maximal tori exist in K.

(iii) If $\dim(K) > 0$, then nontrivial tori exist in K.

Proof. (i) We will verify the equivalence in contrapositive form.

If S is a torus in K that strictly contains T, then (by the Lie correspondence) $\mathfrak{s}$ is an abelian subalgebra of $\mathfrak{k}$ that strictly contains $\mathfrak{t}$.

Conversely, suppose $\mathfrak{s}$ is an abelian subalgebra of $\mathfrak{k}$ that strictly contains $\mathfrak{t}$. Let $x \in \mathfrak{s} - \mathfrak{t}$. Then $\exp(\mathbb{R}x)$ commutes with T, so $\exp(\mathbb{R}x)T$ is an abelian subgroup of G. Its closure $\overline{\exp(\mathbb{R}x)T}$ is a closed abelian subgroup of G, hence compact. The connected component $T' := (\overline{\exp(\mathbb{R}x)T})^0$ of that closure is

then connected, compact and abelian, hence is a torus. Since $x \in \mathrm{Lie}(T') - \mathfrak{t}$, we see that T' strictly contains T.

(ii) Use that the dimensions of the (abelian) subalgebras of $\mathfrak{k}$ are bounded from above.

(iii) Consider $\overline{(\exp(\mathbb{R}x))}^0$ for any $0 \neq x \in \mathfrak{k}$.

$\square$

$$x + y + \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

7.3. Roots and root space decomposition.

Let K be a compact Lie group with maximal torus T. Let G denote the complexification of K, so that $\mathfrak{g}$ is the complexification of $\mathfrak{k}$. We then have the (complexified) adjoint representation

$$\mathrm{Ad} : K \to \mathrm{GL}(\mathfrak{g}),$$

which we may restrict to T and decompose:

$$\mathfrak{g} = \oplus_{\lambda \in \mathfrak{t}_{\mathbb{Z}}^*} \mathfrak{g}^\lambda. \tag{42}$$

Definition 90. A *root* for (K, T) is a nonzero element α of $\mathfrak{t}_{\mathbb{Z}}^*$ for which $\mathfrak{g}^\alpha$ is nonzero. We denote by $\Phi := \Phi(K : T)$ the set of roots, and refer to $\mathfrak{g}^\alpha$ as the *root space* attached to the root α.

For instance, suppose that $K = \mathrm{U}(n)$, with $T = \mathrm{U}(1)^n \hookrightarrow K$ the diagonal subgroup. We've seen (§3.3) that Φ consists of the set of differences $\varepsilon_i - \varepsilon_j$, where i, j are distinct elements of $\{1..n\}$ and ε_i denote the standard basis elements of $\mathfrak{t}_{\mathbb{Z}}^*$. The root space $\mathfrak{g}^{\varepsilon_i - \varepsilon_j}$ is spanned by the elementary matrix E_{ij}, for which $\mathrm{Ad}(t)E_{ij} = (t_i/t_j)E_{ij} = t^{\varepsilon_i - \varepsilon_j} E_{ij}$. For instance, if $n = 3$, then

$$\mathrm{Ad}\begin{pmatrix} t_1 & & \\ & t_2 & \\ & & t_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & (t_1/t_2)a_{12} & (t_1/t_3)a_{13} \\ (t_2/t_1)a_{21} & a_{22} & (t_2/t_3)a_{23} \\ (t_3/t_1)a_{31} & (t_3/t_2)a_{32} & a_{33} \end{pmatrix}.$$

We might observe in this case that $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{C})$ is the direct sum of the diagonal subspace $\mathfrak{t}_{\mathbb{C}}$ and the root spaces, each of which is one-dimensional. We'll see eventually that these features are general, i.e., hold for any compact Lie group. For starters:

Lemma 91. *Let K be a compact Lie group with maximal torus T and $\mathfrak{g}$ as usual. Then*

$$\mathfrak{g} = \mathfrak{t}_{\mathbb{C}} \oplus (\oplus_{\alpha \in \Phi} \mathfrak{g}^\alpha). \tag{43}$$

Proof. By (42), it suffices to show that $\mathfrak{g}^0 = \{x \in \mathfrak{g} : \mathrm{Ad}(t)x = x \text{ for all } t \in T\}$ is equal to $\mathfrak{t}_{\mathbb{C}}$. We note that $\mathfrak{g}^0$ is closed under complex conjugation, hence $\mathfrak{g}^0 = \mathfrak{k}^0 \oplus i\mathfrak{k}^0$ with $\mathfrak{k}^0 := \{x \in \mathfrak{k} : \mathrm{Ad}(t)x = x \text{ for all } t \in T\}$; indeed if $x \in \mathfrak{g}$, then we may write $x = x_1 + ix_2$ with $x_1, x_2 \in \mathfrak{k}$, while for $t \in T$ we have $\mathrm{Ad}(t)x = \mathrm{Ad}(t)x_1 + i\,\mathrm{Ad}(t)x_2$, hence $x \in \mathfrak{g}^0$ iff $x_1, x_2 \in \mathfrak{k}^0$, which leads to the required decomposition. It is thus enough to show that $\mathfrak{k}^0 = \mathfrak{t}$. Suppose otherwise that there exists $x \in \mathfrak{k}^0 - \mathfrak{t}$. Then $\mathbb{R}x \oplus \mathfrak{t}$ is an abelian subalgebra of $\mathfrak{k}$ (note that $[x, \mathfrak{t}] = \{0\}$ because $\mathrm{Ad}(t)x = x$ for all $t \in T$) that strictly contains $\mathfrak{t}$. Since T is a maximal torus, we obtain a contradiction via Lemma 89. $\square$

The root spaces of $\mathfrak{g}$ move the weight spaces of any other representation V predictably, as we now explain. Let K be a compact Lie group with maximal torus T and complexification G. Let $\pi : K \to \mathrm{GL}(V)$ be a finite-dimensional representation of K. It induces a representation $d\pi : \mathfrak{k} \to \mathrm{End}(V)$ of the Lie algebra, which complexifies to a holomorphic representation $d\pi : \mathfrak{g} \to \mathrm{End}(V)$ of the complexified Lie algbera. The basic relationship between these is that for $g \in K$, $x \in \mathfrak{g}$ and $v \in V$, we have

$$\pi(g)d\pi(x)v = d\pi(\mathrm{Ad}(g)x)\pi(g)v.$$

Abbreviating $gv := \pi(g)v$ and $xv := d\pi(x)v$ and $gxg^{-1} := \mathrm{Ad}(g)x$, this identity reads more simply as

$$gxv = (gxg^{-1})gv.$$

Lemma 92. *For $\alpha, \lambda \in \mathfrak{t}_{\mathbb{Z}}^*$, we have*

$$d\pi(\mathfrak{g}^\alpha)V^\lambda \subseteq V^{\lambda+\alpha}.$$

In particular, taking $(\pi, V) = (\mathrm{Ad}, \mathfrak{g})$, we have for any $\alpha, \beta \in \mathfrak{t}_{\mathbb{Z}}^*$ that

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subseteq \mathfrak{g}^{\alpha+\beta}.$$

Proof. Let $x \in \mathfrak{g}^*$, $v \in V^\lambda, t \in T$. Then

$$\pi(t)(d\pi(x)v) = d\pi(\mathrm{Ad}(t)x)\pi(t)v = d\pi(t^\alpha x)\pi(t)v = d\pi(t^\alpha x)t^\lambda v = t^{\lambda+\alpha}d\pi(x)v, \tag{44}$$

so $d\pi(x)v \in V^{\lambda+\alpha}$, as required. We record the same argument but expressed using the above abbreviations:

$$txv = (txt^{-1})tv = (t^\alpha x)(t^\lambda v) = t^{\lambda+\alpha}xv.$$

The conclusion regarding α, β is immediate, but it's worth repeating the proof in that special case for the sake of illustration: for $x \in \mathfrak{g}^\alpha$, $y \in \mathfrak{g}^\beta$ and $t \in T$, we have

$$\mathrm{Ad}(t)[x,y] = [\mathrm{Ad}(t)x, \mathrm{Ad}(t)y] = [t^\alpha x, t^\beta y] = t^{\alpha+\beta}[x,y],$$

hence $[x,y] \in \mathfrak{g}^{\alpha+\beta}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

[12]

By choosing a unitary structure on a faithful representation of K and then taking an orthonormal basis consisting of weight vectors for T, we obtain an embedding $K \hookrightarrow \mathrm{U}(n)$ that carries T to a subgroup of the diagonal subgroup of $\mathrm{U}(n)$. We have noted (Theorem 75) that then the complexified Lie algebra $\mathfrak{g}$ is closed under $\theta : x \mapsto -x^*$, where $x^* := {}^t\overline{x}$. This fact has an important consequence:

Lemma 93. *If $\alpha \in \Phi$, then $-\alpha \in \Phi$, and $\dim(\mathfrak{g}^\alpha) = \dim(\mathfrak{g}^{-\alpha})$.*

For example, if $K = \mathrm{U}(n)$ and (with notation as before) $\alpha = \varepsilon_i - \varepsilon_j$, then $\mathfrak{g}^\alpha = \mathbb{C}E_{ij}$, $\theta(E_{ij}) = -E_{ji}$, $-\alpha = \varepsilon_j - \varepsilon_i$.

Proof. Since $\mathfrak{g}$ is closed under θ, it is enough to show that $\theta(\mathfrak{g}^\alpha) \subseteq \mathfrak{g}^{-\alpha}$. (Equality then follows from, for instance, the involutivity of θ and the same argument applied to $-\alpha$.) Let $x \in \mathfrak{g}^\alpha$, and let $t \in T$. With matrix realizations as above, we have

---

[12]End of lecture #11, Thursday, 28 March

$t^* = t^{-1}$ (indeed, we may write $t = \mathrm{diag}(e^{i\theta_1}, \ldots e^{i\theta_n})$, which makes this obvious), while $A^*B^* = (BA)^*$ for any matrices A, B, thus

$$\mathrm{Ad}(t)\theta(x) = -tx^*t^{-1} = -(t^{-1})^*x^*t^*$$
$$= -(txt^{-1})^* = -(\mathrm{Ad}(t)x)^* = -(t^\alpha x)^* = \overline{t^\alpha}\theta(x) = t^{-\alpha}\theta(x).$$

Since t was arbitrary, we deduce as required that $\theta(x) \in \mathfrak{g}^{-\alpha}$. $\qquad\square$

7.4. Generators. Tori are topologically cyclic, and this fact is very useful.

Definition 94. Let T be any torus. A *generator* $t \in T$ is an element such that $\overline{\langle t \rangle} = T$; here $\langle t \rangle := \{t^n : n \in \mathbb{Z}\}$ denotes the subgroup generated by t, and the closure of that subgroup $\overline{\langle t \rangle}$.

Lemma 95. *Let T be any torus. Then generators of T exist and are dense.*

Proof. Write $n = \dim(T)$, so that $T \cong (\mathbb{R}/\mathbb{Z})^n$ and $t = (x_1, \ldots, x_n)$. In the case $n = 1$, Weyl's equidistribution criterion implies that t is a generator iff $x_1 \notin \mathbb{Q}$. For general n, the same criterion says that t is a generator iff $1, x_1, \ldots, x_n$ are linearly independent over $\mathbb{Q}$. These conditions are obviously satisfied by a dense collection of elements.

Alternatively, here's an elementary pigeonhole argment (which I think I learned from Adams' book). For each $t \in T$ and $\varepsilon > 0$, write $B_t(\varepsilon)$ for the open ε-ball at t, defined with respect to the Euclidean metric on $(\mathbb{R}/\mathbb{Z})^n$. Fix an enumeration (t_i, ε_i), indexed by $i \geq 1$, of all pairs consisting of

- an element $t_i \in T$ whose entries with respect to the isomorphism $T \cong (\mathbb{R}/\mathbb{Z})^n$ are rational, and
- a rational number $\varepsilon \in (0, 1)$.

(Thus the $B_{t_i}(\varepsilon_i)$ give a countable open basis for T.) Let U be a closed subset of T with nonempty interior U^0. If the natural number N_1 is large enough, then $N_1 U := \{N_1 u : u \in U\}$ coincides with T. Thus $U_1 := \{u \in U : |N_1 u - t_1| \leq \varepsilon_1/2\}$ is a closed subset of U, with nonempty interior, such that $N_1 U_1 \subseteq B_{t_1}(\varepsilon_1)$. We may similarly find a natural number N_2 and a closed subset U_2 of U_1, with nonempty interior, so that $N_2 U_2 \subseteq B_{t_2}(\varepsilon_2)$. Continuing in this way, we obtain a descending chain of closed subsets with nonempty interiors $U \supset U_1 \supseteq U_2 \supseteq \cdots$ and natural numbers $N_1, N_2, \ldots$ so that for each $i \geq 1$, we have $N_i U_i \subseteq B_{t_i}(\varepsilon_i)$. By compactness, the intersection $\cap_{i \geq 1} U_i$ is nonempty. Let t be any element of that intersection. Then $N_i t \in B_{t_i}(\varepsilon_i)$. Hence the group $\langle t \rangle = \mathbb{Z}t$ (here we use additive notation for $T \cong (\mathbb{R}/\mathbb{Z})^n$) intersects every constituent $B_{t_i}(\varepsilon_i)$ of an open basis of T. Hence its closure $\overline{\langle t \rangle}$ is all of T. Since U was arbitrary, we deduce moreover that the set of generators is dense. $\qquad\square$

Generators are useful because they reduce properties involving an entire torus to individual elements. We give an example. To state it, recall that the centralizer and normalizer of a subgroup H of a group G are defined by $Z_G(H) := \{g \in G : ghg^{-1} = h \text{ for all } h \in H\}$ and $N_G(H) := \{g \in G : ghg^{-1} \in H \text{ for all } h \in H\}$; the definition of $Z_G(H)$ applies more generally to any subset H of G, and we abbreviate $Z_G(h) := Z_G(\{h\})$ when that subset consists of a single element $h \in G$.

Lemma 96. *Let K be a Lie group, let $T \leq K$ be any torus, and let $t \in T$ be a generator. Then $Z_K(T) = Z_K(t) = \{g \in K : gtg^{-1} = t\}$ and $N_K(T) = \{g \in K : gtg^{-1} \in T\}$*

Proof. We just note that the subsets $\{s \in T : gsg^{-1} = s\}$ and $\{s \in T : gsg^{-1} \in T\}$ are closed subgroups of T, hence they coincide with T iff they contain the generator t. $\qquad\square$

7.5. Definition and finitude of the Weyl group of a maximal torus.

Definition 97. Let K be a compact Lie group, with maximal torus T. The *Weyl group* $W := W(K : T)$ is defined to be $N(T)/T$, where $N(T) := N_K(T)$ denotes the normalizer of the torus.

For instance, for $K = \mathrm{U}(n)$, we saw that the permutation representation $S(n) \to \mathrm{U}(n)$ defines an isomorphism $\S(n) \cong W$.

We note in general that W acts on T by conjugation (in a well-defined manner):
$$w \cdot t := wtw^{-1}.$$

The following basic lemma (which may be proved in a few ways) gives a decent illustration of how to argue using the root space decomposition (43).

Lemma 98. Let K be a compact Lie group with maximal torus T. Then $N(T)^0 = T$ and $|W| < \infty$.

Proof. We show first that $N(T)^0 = T$. The containment $N(T)^0 \supseteq T$ is clear, so we just need to check that $\mathrm{Lie}(N(T)^0) \subseteq \mathfrak{t}$. We have $\mathrm{Lie}(N(T)^0) = \{x \in \mathfrak{k} : [x, \mathfrak{t}] \subseteq \mathfrak{t}\}$. Let $x \in \mathrm{Lie}(N(T)^0)$. Using (43), we may write
$$x = x_0 + \sum_{\alpha \in \Phi} x_\alpha$$
with $x_0 \in \mathfrak{t}_{\mathbb{C}}$ and $x_\alpha \in \mathfrak{g}^\alpha$. For $y \in \mathfrak{t}$, we then have
$$\mathfrak{t} \ni -[x, y] = [y, x] = [y, x_0] + \sum_{\alpha \in \Phi} [y, x_\alpha] = \sum_{\alpha \in \Phi} \alpha(y) x_\alpha,$$
hence $\alpha(y) x_\alpha = 0$ for all α and all y. Since each α is nonzero, we deduce that $x_\alpha = 0$ for all α, hence that $x = x_0$, and so $x \in \mathfrak{t}$, as required.

It follows that $W = N(T)/N(T)^0$ is a compact 0-dimensional Lie group, and so is finite. $\qquad\square$

Remark 99. The definition of W given here is a bit *ad hoc*. We'll later define $W(K : S)$, for *any* torus S, to be the quotient $N_K(S)/Z_K(S)$ of the normalizer divided by the centralizer. This will be seen to coincide with the definition given above in the case of a *maximal* torus T only after we've seen that $Z_K(T) = T$ for such tori, but this is a deep fact which will require proof. The group $W = W(K : T)$ as defined above will be relevant for the proof of that fact.

7.6. Weyl integral formula.
We want a generalization of the formula that we proved earlier for $\mathrm{U}(n)$ (see §3.4). That formula expressed the integral over K in terms of integrals over conjugates of a maximal torus T. Since tori are connected, we'd better assume that K is connected to have any hope for such a formula.

So let K be a connected compact Lie group, with maximal torus T. Write $N := \dim(K)$ and $n := \dim(T)$. The basic setup will be the same as in the case of $\mathrm{U}(n)$. Write $N := \dim(K)$ and $n := \dim(T)$. The subspace $\mathfrak{t} = \mathrm{Lie}(T)$ of $\mathfrak{k} = \mathrm{Lie}(K)$ admits a natural $\mathrm{Ad}(T)$-invariant complement $\mathfrak{k}/\mathfrak{t} := \mathfrak{k} \cap (\sum_{\alpha \in \Phi} \mathfrak{g}^\alpha)$. Fix a nonzero $\omega_1 = \alpha_1 \wedge \beta_1$ in $\Lambda^N(\mathfrak{k})$, with $\alpha_1 \in \Lambda^n(\mathfrak{t})$ and $\beta_1 \in \Lambda^{N-n}(\mathfrak{k}/\mathfrak{t})$. The compactness of T implies that β_1 is $\mathrm{Ad}(T)$-invariant. We obtain left-invariant

differential forms $\omega \in \Omega^N(K), \alpha \in \Omega^n(T), \beta \in \Omega^{N-n}(K/T)$. (We'll denote here and henceforth by $\Omega^k(M)$ for the space of smooth k-forms on a manifold M, and $\Omega_c^k(M)$ for the subspace of compactly-supported k-forms.) The compactness of K implies that ω is also right-invariant. We obtain Haar measures $|\omega|, |\alpha|, |\beta|$ on $K, T, K/T$. We may assume ω and α normalized so that $|\omega|$ and $|\alpha|$ are probability Haar measures; the same is then true for $|\beta|$ by the analogue of Fubini's theorem. We define

$$q : K/T \times T \to K$$
$$q(g,t) := gtg^{-1},$$

and write

$$q^*\omega = \det(q)(\beta \wedge \alpha).$$

Lemma 100.   *(i) For all $(g,t) \in K/T \times T$, we have*

$$\det(q)(g,t) = D(t),$$

where

$$D(t) := \det(1 - \mathrm{Ad}(t)|\mathfrak{k}/\mathfrak{t}) = \prod_{\alpha \in \Phi}(1 - t^\alpha)^{\dim(\mathfrak{g}^\alpha)}.$$

(ii) We have $D(t) \geq 0$ for all $t \in T$. We have $D(t) > 0$ precisely when $t^\alpha \neq 1$ for all $\alpha \in \Phi$.

We'll see a bit later that each root space $\mathfrak{g}^\alpha$ is exactly one-dimensional, so that the dimensions appearing in the exponents may be omitted, but we include them for now.

Proof. The computation of $\det(q)$ is exactly as in the case of $\mathrm{U}(n)$.

By lemma 93, we have $\alpha \in \Phi$ iff $-\alpha \in \Phi$, and moreover $\dim \mathfrak{g}^\alpha = \dim \mathfrak{g}^{-\alpha}$. Thus $D(t)$ is a product taken over all distinct pairs $\{\alpha, -\alpha\} \subseteq \Phi$ of the $\dim(\mathfrak{g}^\alpha)$th power of the quantities $(1 - t^\alpha)(1 - t^{-\alpha})$. Each such quantity is nonnegative. (Writing $t^\alpha = e^{i\theta}$, the quantity in question is $2 - 2\cos\theta$.) $\square$

Lemma 101.   *The formula*

$$D(g) := \text{coefficient of } \lambda^{\dim(T)} \text{ in } \det(\lambda + 1 - \mathrm{Ad}(g))$$

defines a class function $D : K \to \mathbb{R}$ that extends the function $D : T \to \mathbb{R}_{\geq 0}$ defined above.

We'll see eventually that in fact $D \geq 0$ on all of K.

Proof. For $t \in T$, $\mathrm{Ad}(t)$ acts on both $\mathfrak{t}$ and $\mathfrak{k}/\mathfrak{t}$, so we may factor the determinant of $\lambda + 1 - \mathrm{Ad}(t)$ for the action on on $\mathfrak{k}$ as the product of the determinants for the actions on $\mathfrak{t}$ and $\mathfrak{k}/\mathfrak{t}$; we have $\mathrm{Ad}(t)|_\mathfrak{t} = 1$, so the first of the latter two determinants is just $\lambda^{\dim(T)}$, while the second is $D(t) + \mathrm{O}(\lambda)$. $\square$

Definition 102. An element $g \in K$ is called *regular* if $D(g) \neq 0$. We denote by K^{reg} the set of all regular elements $g \in K$, and set $T^{\mathrm{reg}} := T \cap K^{\mathrm{reg}}$.

We observe that K^{reg} is a conjugacy-invariant subset of K and that $T^{\mathrm{reg}} = \{t \in T : t^\alpha \neq 1 \text{ for all } \alpha \in \Phi\}$. For example, if $K = \mathrm{U}(n)$ and $T \cong \mathrm{U}(1)^n$ is the diagonal torus, then $t \in T$ is regular iff $t_i \neq t_j$ for all $i \neq j$. It's clear that any generator is regular, while the converse is not true in general. Note also that "being a generator" is intrinsic to T, while "being regular" depends also upon the ambient group K.

We note that for any $w \in W$ and $(gT, t) \in K/T \times T$ and $w \in W$, the action

$$w \cdot (gT, t) := (gw^{-1}T, w \cdot t) = (gw^{-1}T, wtw^{-1})$$

is well-defined. It induces defines an action of W on each fiber of q; indeed,

$$q(w \cdot (gT, t)) = (gw^{-1})(wtw^{-1})(gw^{-1})^{-1} = gtg^{-1} = q(gT, t).$$

We denote this relationship by the diagram

$$W \circlearrowright K/T \times T \xrightarrow{q} K.$$

In the $\mathrm{U}(n)$ case, the fibers of q correspond to choices of orthonormal basis for a given unitary matrix; the action of W corresponds to permuting the basis elements and is thus free, and transitive on the fiber above a regular element because such elements admit a unique (up-to-reordering) basis of eigenvectors.

The action of W (on each fiber of q) is free in general, since if $(gwT, w \cdot t) = (gT, t)$ then $gwT = gT$, hence $wT = T$, i.e., $w \equiv 1$ in W. We'll see below that the action on regular fibers is transitive, as in the $\mathrm{U}(n)$ case. The action can be far from transitive on irregular fibers; for instance, $q^{-1}(1) = K/T \times \{1\}$, and the action of W on K/T fails to be transitive whenever $\dim(K/T) > 0$.

We're now prepared to state the Weyl integral formula for K, together with some related consequences:

Theorem 103. *Let notation and assumptions be as above: K is an N-dimensional connected compact Lie group with n-dimensional maximal torus T and Weyl group W.*

(i) q is surjective, i.e., $K = \cup_{g \in K} gTg^{-1}$.

(ii) The restriction

$$q : K/T \times T^{\mathrm{reg}} \to K^{\mathrm{reg}}$$

is a $|W|$-to-one covering map, with the group W acting simply-transitively on the fibers. This restriction is orientation-preserving at every point.

(iii) For all $\nu \in \Omega^N(K)$, we have

$$\int_{K/T \times T} q^* \nu = |W| \int_K \nu. \tag{45}$$

(iv) For all $f \in C(K)$ (i.e., continuous $f : K \to \mathbb{C}$) we have

$$\int_{(g,t) \in K/T \times T} D(t) f(gtg^{-1}) = |W| \int_K f,$$

with all integrals taken with respect to probability Haar measures.

Proof that iii implies iv. Take $\nu = f\omega$, with $\omega \in \Omega^N(K)$ as used above to describe the probability Haar on K. Then $q^* \nu = f q^* \omega = f D(\beta \wedge \alpha)$. □

Proof that iii implies i. Let E denote the image of q. Suppose otherwise that we may find some $x \in K - E$. Since $K/T \times T$ is compact and q is continuous, the set E is compact, hence closed, so we may find an open neighborhood U of x disjoint from E. Choose $\nu \in \Omega_c^N(U)$ with $\int \nu \neq 0$. Then $q^* \nu = 0$. Using (ii), we deduce that

$$0 = \int q^* \nu = |W| \int \nu \neq 0,$$

giving the required contradiction. □

Proof of ii and iii. First, let $x \in T^{\mathrm{reg}}$. Lemma 100 implies then that $\det(q) > 0$ on $q^{-1}(x)$. Thus for each $y \in q^{-1}(x)$, the map q is an orientation-preserving local diffeomorphism at y. In particular, $q^{-1}(x)$ is a closed 0-dimensional subset of the compact set $K/T \times T$, and so $q^{-1}(x)$ is finite. We may thus find a small neighborhood U of x and, for each $y \in q^{-1}(x)$, a small neighborhood U_y of y so that q induces diffeomorphisms $U_y \cong U$. By shrinking U and hence each U_y if necessary, we may assume that the U_y are all disjoint. Using the compactness of the domain of q, we see that the image under q of the complement of $\cup_z U_z$ is compact, hence closed, so by shrinking U if necessary, we may assume that the image in question is disjoint from U. Then

$$q^{-1}(U) = \sqcup_{y \in q^{-1}(x)} U_y,$$

with the map $q^{-1}(U) \to U$ a trivial cover.

We've noted already that W acts freely on every fiber of q. We claim now moreover that W acts *transitively* on every fiber of q in U, i.e., that W permutes the "pancakes" U_y lying above U. It suffices to verify that W acts transitively on the fibers above any individual element $t \in U$. By Lemma 95, we may take for t a generator of T. As a "basepoint" for $q^{-1}(t)$, we may take the pair (eT, t), with $e \in G$ the identity element. If (gT, s) is any other element of $q^{-1}(t)$, then $gsg^{-1} = t$, and so $g^{-1}tg = s \in T$, which (by Lemma 96) forces g to lie in $N(T)$, i.e., $w := gT$ to lie in W. Thus $(gT, s) = w \cdot (eT, t)$, and so W acts transitively on $q^{-1}(t)$, as required. It follows in particular that $\#q^{-1}(x) = \#q^{-1}(t) = |W|$, so that for any $\nu \in \Omega_c^N(U)$,

$$\int q^*\nu = |W| \int \nu. \tag{46}$$

We have nearly established ii; what remains to show is just that $q : K/T \times T^{\mathrm{reg}} \to K^{\mathrm{reg}}$ is surjective. The required surjectivity will follow from i and hence from iii, so it remains only to establish iii.

To that end, let us call an open subset U of K *good* if (45) holds for all $\nu \in \Omega_c^N(U)$. By a partition of unity argument and the compactness of K, it will suffice to show that every element of K admits a good neighborhood. We've seen already in (46) that every element of T^{reg} admits a good neighborhood. The set T^{reg} is nonempty. The idea now will be to use Stokes' theorem to "propagate" goodness from any one open set to the rest. The following elementary fact from differential geometry, known as the *Poincaré lemma*, will be of use:

Let $\nu_1, \nu_2 \in \Omega_c^N((0,1)^N)$ be compactly-supported top-degree differential forms on the open unit cube such that $\int \nu_1 = \int \nu_2$. Then there exists $\lambda \in \Omega_c^{N-1}((0,1)^N)$ so that $\nu_2 = \nu_1 + d\lambda$.

For example, if $N = 1$, then $\nu_1 - \nu_2 = f(x)dx$ for some $f \in C_c^\infty((0,1))$ with $\int_0^1 f(x)\,dx = 0$, so $\lambda(x) := \int_0^x f(y)\,dy$ defines an element $\lambda \in C_c^\infty((0,1))$ for which $d\lambda = \nu_1 - \nu_2$. The proof in the general case is similar and recorded below for completeness. Assuming the Poincaré lemma for the moment, we may complete the proof of iii as follows:

(1) Say that an open subset U of K is *small* if it admits a chart $U \cong (0,1)^N$. We claim that *if U_1, U_2 are small, $U_1 \cap U_2$ is nonempty, and U_1 is good, then U_2 is good.* To see this, let $\nu_2 \in \Omega_c^N(U_2)$ be given. Choose any $\nu_1 \in \Omega_c^N(U_1 \cap U_2)$ with $\int \nu_1 = \int \nu_2$. By the Poincaré lemma applied in the chart for U_2, we may find $\lambda \in \Omega_c^{N-1}(U_2)$ so that $\nu_2 = \nu_1 + d\lambda$. We

have $q^* d\lambda = dq^* \lambda$, whose integral vanishes thanks to Stokes' theorem. By the assumed goodness of U_1 and the construction of ν_1, we deduce that $\int q^* \nu_2 = \int q^* \nu_1 = |W| \int \nu_1 = |W| \int \nu_2$. Since ν_2 was arbitrary, we deduce as required that U_2 is good.

(2) We have seen that there exists *some* small good subset U of K (namely, any small enough open neighborhood of a regular element of T). It will suffice to show that any other small open subset U' of K is likewise good. To that end, let us choose a curve $\gamma : [0,1] \to K$ with $\gamma(0) \in U$ and $\gamma(1) \in U'$. We can find for each $s \in [0,1]$ a good neighborhood of $\gamma(s)$. By a compactness argument applied to the preimages of those neighborhoods, we may find a sequence $U = U_0, U_1, \ldots, U_{m-1}, U_m = U'$ of small open subsets of K with $U_j \cap U_{j-1} \neq \emptyset$ for all $j = 1..m$. By what was shown above, the goodness of $U_0 = U$ inductively implies the goodness of every U_j, hence in particular of U'.

The proof is now complete. $\qquad\square$

Proof of the Poincaré lemma. (Omitted from lecture, recorded here for completeness; probably a good exercise.) Let's work with n instead of N. Write

$$\nu_1 - \nu_2 = f(x_1, \ldots, x_n)\, dx_1 \wedge \cdots \wedge dx_n.$$

By hypothesis, f is supported in $(0,1)^n$ and has mean zero. Our aim is to show that $f(x_1, \ldots, x_n)\, dx_1 \wedge \cdots \wedge dx_n$ is of the form $d\lambda$ for some $(n-1)$-form λ supported in $(0,1)^n$. Equivalently, we must show that

$$f = \sum_{j=1}^{n} \partial_j g_j \tag{47}$$

for some $g_1, \ldots, g_n$ supported in $(0,1)^n$, where ∂_j denotes the partial derivative taken with respect to x_j.

We induct on n. The case $n = 0$ is tautological, while the case $n = 1$ was treated above. For the case $n \geq 2$, we apply our inductive hypothesis to write

$$\int_{y=0}^{1} f(y, x_2, \ldots, x_n)\, dy = \sum_{j=2..n} \partial_j h_j(x_2, \ldots, x_n).$$

Fixing $\chi \in C_c^\infty((0,1))$ with $\int \chi = 1$, the function

$$g_j(x_1, \ldots, x_n) := \chi(x_1) h_j(x_2, \ldots, x_n)$$

is supported in $(0,1)^n$. Set $f' := f - \sum_{j \geq 2} \partial_j g_j$. Then

$$\int_{y=0}^{1} f'(y, x_2, \ldots, x_n)\, dy = 0 \text{ for all } x_2, \ldots, x_n,$$

so that

$$g_1(x_1, \ldots, x_n) := \int_{y=0}^{x_1} f'(y, x_2, \ldots, x_n)\, dy$$

is supported in $(0,1)^n$ and satisfies $\partial_1 g_1 = f'$. The required identity (47) then holds by construction. $\qquad\square$

Remark 104. I wanted to present the proof as elementarily as possible, but I would be remiss not to mention that what the above arguments "really" show

is that for any connected N-manifold M, the compactly-supported top-degree de Rham cohomology group

$$H_c^N(M) := \Omega_c^N(M)/d\Omega_c^{N-1}(M),$$

with coefficients taken in either $k = \mathbb{R}$ or $\mathbb{C}$, is isomorphic to k, with the isomorphism

$$H_c^N(M) \xrightarrow{\cong} k \tag{48}$$

defined by integrating an N-form over M. (Convince yourself that our arguments actually establish this.) Since $K/T \times T$ and K are both connected compact N-manifolds, we can replace H_c^N with H^N, and the formula (45) says that the dual map q^* on de Rham cohomlogy fits into a commutative diagram

$$
\begin{array}{ccc}
H^N(K) & \xrightarrow{\ \ q^*\ \ } & H^N(K/T \times T) \\
{\scriptstyle\cong}\big\downarrow & & \big\downarrow{\scriptstyle\cong} \\
k & \xrightarrow[\ |W|\ ]{} & k,
\end{array}
$$

where we write $|W|$ for the "multiplication by $|W|$" map. We know that every continuous k-linear map $k \to k$ is given by multiplication by *something*, so the issue is just to compute the proportionality constant for at least one N-form ν. We managed to do this for all ν supported in a small enough neighborhood of a regular element of the torus.

It's also worth noting that one can see *a priori* that proportionality constant (in this case $|W|$) must an *integer* by applying the comparison isomorphism relating de Rham cohomology to singular cohomology and working instead with $\mathbb{Z}$-coefficients for the latter, using Poincaré duality to obtain the analogue over $\mathbb{Z}$ of the integration isomorphism (48). For details see the BTD course reference.

Remark 105. The intuitive idea behind the proof of the surjectivity assertion i of Theorem 103 may be illustrated as follows. Let C be a smooth periodic oriented curve in the "cylinder" $X := \mathbb{R}/\mathbb{Z} \times \mathbb{R}$, thus C is the image of some smooth map $\gamma : \mathbb{R}/\mathbb{Z} \to X$. Consider the map $\phi : C \to \mathbb{R}/\mathbb{Z}$ given by the projection $(x, y) \mapsto x$ onto the first coordinate. Suppose that ϕ fails to be surjective; say $x_0 \notin \phi(C)$. Take a "generic" nonempty fiber $\phi^{-1}(x)$. It's then intuitively plausible that the signed sum of the orientations of the elements $y \in \phi^{-1}(x)$ must be zero, because the trajectories along C departing from such elements and moving in the direction of x_0 must pair off with one another in a complementary fashion to be compatible with $\phi^{-1}(x_0)$ being empty. This intuitive plausibility is in fact the case, and may be formulated and proved rigorously by the above arguments. This explains why q failing to be surjective is incompatible with the fact that $\det(q)$ is everywhere positive, the latter of which implies that the orientations sum to $|W|$ rather than to 0.

13

-----

[13]End of lecture #12, Tuesday, 2 Apr

7.7. **Consequences of the conjugacy theorem.** Let K be a compact connected Lie group. We saw above that for any maximal torus T, we have $K = \cup_{g \in K} gTg^{-1}$. We now explore some consequences of this fact.

Theorem 106. *Every element of K is contained in some maximal torus, the exponential map $\exp : \mathfrak{k} \to K$ is onto, and every $x \in K$ is contained in the connected component $Z_K(x)^0$ of its centralizer.*

Proof. If T is a maximal torus, then its conjugates gTg^{-1} are also maximal tori, and we've seen that every element of K belongs to one of these. Since for any torus T the exponential map $\exp : \mathfrak{t} \to T$ is surjective, we deduce the corresponding result for K. The final assertion is clear when x is contained in the maximal torus T, because then $g \in T \leq Z_K(x)^0$, and follow in general from what we have just shown. $\qquad\square$

It's worth illustrating the subtlety here with a "near proof." Given an element $x \in K$, let $H := \overline{\langle x \rangle}$ denote the closure of the subgroup that it generates. Then H is closed in K, hence compact, and abelian. If we knew that H were connected, then H would be a torus, and so we'd be done. If we even knew merely that x belonged to the connected component H^0 of H, then we could conclude similarly. But in general $x \notin H^0$. For instance, take $x = (-1, e^i) \in \mathrm{U}(1)^2$. Then $H = \{\pm 1\} \times \mathrm{U}(1)$ and $H^0 = \{1\} \times \mathrm{U}(1)$, so $x \notin H^0$. This illustrates some of the subtlety involved in Theorem 106.

Anyway, continuing with the consequences:

Theorem 107. *Any two maximal tori T, T' of K are conjugate.*

Proof. Let $t' \in T'$ be a generator. We've seen that $t' \in gTg^{-1}$ for some $g \in K$. Since gTg^{-1} is a closed subgroup and t' is a generator, it follows that $T' \subseteq gTg^{-1}$. Since T is a maximal torus, so is gTg^{-1}, which forces the equality $T' = gTg^{-1}$. $\quad\square$

Theorem 108. *The center $Z(K)$ of K is given by*
$$Z(K) = \cap_{\text{maximal tori } T} T.$$

Proof.
- "$\supseteq$": Let $z \in \cap T$ and $x \in K$. By Theorem 106, we may find a maximal torus T with $x \in T$. Then $z \in T$, so x and z commute. Since x was arbitrary, we conclude that $z \in Z(K)$.
- "$\subseteq$": Let $z \in Z(K)$, and let T be a maximal torus. Choose $g \in K$ so that $z \in gTg^{-1}$. Then $T \ni g^{-1}zg = z$, because z is central. Since T was arbitrary, we conclude that $z \in \cap T$.

$\square$

Theorem 109. *Let T be a maximal torus, with Weyl group $W = N(T)/T$. If $t, t' \in T$ are K-conjugate, then they have the same W-orbits. Thus the map*
$$T/W \to \{\text{conjugacy classes in } K\}$$
is bijective.

Proof. Let $H := Z_K(t')^0$ denote the connected component of the centralizer of $t' \in T$. Choose $g \in K$ so that $t' = gtg^{-1}$; in paricular, $t' \in gTg^{-1}$. Then the tori gTg^{-1} and T are both contained in H and maximal in K, hence maximal in H, so by Theorem 107 applied to the compact connected Lie group H, we see that there is some $h \in H$ so that $T = hgTg^{-1}h^{-1}$, i.e., so that $hg \in N_K(T)$. Let $w := hgT$

denote the class of hg in the Weyl group W. Then $w \cdot t = hgtg^{-1}h^{-1} = ht'h^{-1} = t'$, since h centralizes t'. □

Theorem 110. *Let $S \subseteq K$ be any torus. Then its centralizer is described by the formula*

$$Z_K(S) = \cup_{maximal\ tori\ T \supseteq S} T.$$

Proof. We start by noting that if $g \in Z_K(S)$, then $S \subseteq Z_K(g)^0$. The remainder of the proof is left as an exercise in applying the ideas introduced above. □

Corollary 111. *For all tori S, the centralizer $Z_K(S)$ is connected.*

This might sound at first like a very dry and boring result, but knowing that a subgroup is connected is quite powerful, because it reduces the determination of that subgroup to that of its Lie algebra.

Corollary 112. *Let T be a maximal torus. Then $Z_K(T) = T$. Thus the Weyl group $W = N_K(T)/T = N_K(T)/Z_K(T)$ acts faithfully on T, i.e., W injects into $\mathrm{Aut}(T)$.*

Corollary 113. *Let $x \in \mathfrak{k}$. Then $Z_K(x) := \{g \in K : \mathrm{Ad}(g)x = x\}$ is connected.*

Proof. Set $S := \overline{\exp(\mathbb{R}x)}$. Then S is a torus, and $Z_K(x) = Z_K(S)$. □

7.8. **Basics on Weyl chambers.** We continue to assume that K is a compact connected Lie group. We fix a maximal torus T. We now define the "Lie algebra variants" of the sets of regular elements defined at the group level in §7.6. To that end, let us denote by

$$D_0 : \mathfrak{g} \to \mathbb{R}$$

$$D_0(x) := \text{ coefficient of } \lambda^{\dim(T)} \text{ in } \det(\lambda - \mathrm{ad}(x))$$

the Lie algebra analogue of the map $D : G \to \mathbb{R}$ defined earlier. We may compute D_0 on a toral element $x \in \mathfrak{t}$ as follows. The root space decomposition (43) for the action of T on $\mathfrak{g}$ differentiates to a decomposition for the action of $\mathfrak{t}$ on $\mathfrak{g}$, namely if we decompose $x \in \mathfrak{g}$ as

$$x = x_0 + \sum x_\alpha \in \mathfrak{t}_\mathbb{C} \oplus (\oplus_{\alpha \in \Phi} \mathfrak{g}^\alpha), \tag{49}$$

then

$$[z, x] = 0 + \sum \alpha(z)x_\alpha \text{ for all } z \in \mathfrak{t}. \tag{50}$$

Thus $\det(\lambda - \mathrm{ad}(z)) = \lambda^{\dim(T)} \prod_{\alpha \in \Phi}(\lambda - \alpha(z))^{\dim(\mathfrak{g}^\alpha)}$. Since the roots α come in pairs $\{\alpha, -\alpha\}$ with $\dim(\mathfrak{g}^\alpha) = \dim(\mathfrak{g}^{-\alpha})$, we have $\prod_{\alpha \in \Phi}(-1)^{\dim(\mathfrak{g}^\alpha)} = 1$, which gives the simple formula

$$D_0(z) = \prod_{\alpha \in \Phi} \alpha(z)^{\dim(\mathfrak{g}^\alpha)} \text{ for } z \in \mathfrak{t}.$$

We note again that we'll show in a bit that each root space is one-dimensional, so that the exponents $\dim(\mathfrak{g}^\alpha)$ may be omitted.

Definition 114. We say that $x \in \mathfrak{k}$ is *regular* if $D_0(x) \neq 0$. We write $\mathfrak{k}^{\mathrm{reg}} \subseteq \mathfrak{k}$ for the subset of regular elements. We more generally use a superscripted reg, as in $\mathfrak{t}^{\mathrm{reg}}, \mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$, etc., to denote the subset of regular elements.

A *Weyl chamber* C is a connected component of $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$.

We note that $\mathfrak{k}^{\mathrm{reg}}$ is closed under $\mathrm{Ad}(K)$, since the characteristic polynomial is invariant by conjugation. We leave it to the reader to check the instructive identity

$$\mathfrak{k}^{\mathrm{reg}} = \{x \in \mathfrak{k} : \exp(i\varepsilon x) \in K^{\mathrm{reg}} \text{ for all small } \varepsilon > 0.\}$$

Thus $\mathfrak{k}^{\mathrm{reg}}$ may be regarded as a sort of "tangent cone" at the identity to the open subset K^{reg} of K. Note however that there are elements of $\mathfrak{k}^{\mathrm{reg}}$ whose exponentials do not lie in K^{reg}; for example, $x = \mathrm{diag}(\pi i, -\pi i) \in \mathfrak{su}(2)$ is regular, but $\exp(x) = \mathrm{diag}(-1, -1) \in \mathrm{SU}(2)$ is not regular.

Let C be a Weyl chamber. Its image $\alpha(C)$ under any root α is then a connected subset of $\mathbb{R}_{\neq 0} = \mathbb{R}_{>0} \cup \mathbb{R}_{<0}$, and is thus contaiend in either the positive or negative reals. In other words, we may find signs $\varepsilon_\alpha \in \{\pm 1\}$ ($\alpha \in \Phi$) so that $C \subseteq C_\varepsilon := \{x \in \mathfrak{t}_{\mathbb{R}} : \varepsilon_\alpha \alpha(x) > 0 \text{ for all } \alpha \in \Phi\}$. On the other hand, it's clear from the definition that C_ε is a *convex cone*, i.e., if $x_1, \ldots, x_n \in C_\varepsilon$ and $t_1, \ldots, t_n > 0$, then $t_1 x_1 + \cdots + t_n x_n$. In particular, C_ε is a connected subset of $\mathfrak{t}_{\mathbb{R}}^{\mathrm{reg}}$. Since C is by definition a connected component of the latter set, we deduce that $C = C_\varepsilon$; that is to say, every Weyl chamber is described by requiring that each root α have some prescribed sign ε_α on all of C. In particular, C is a convex cone.

For instance, if $K = \mathrm{U}(n)$, then $\mathfrak{t}_{\mathbb{R}}$ identifies with the set $\mathbb{R}^n$ of n-tuples $x = (x_1, \ldots, x_n)$ of real numbers, and the subset $\mathfrak{t}_{\mathbb{R}}^{\mathrm{reg}}$ with the subset consisting of those x whose entries are distinct: $x_i \neq x_j$ for $i \neq j$. Since the roots are of the form $\varepsilon_i - \varepsilon_j$ with $i \neq j$, any Weyl chamber C is described by inequalities of the form $x_i > x_j$ or $x_i < x_j$ for all indices $i < j$. There is thus a permutation σ of $\{1..n\}$ so that $C = \{x : x_{\sigma(1)} > \cdots > x_{\sigma(n)}\}$. In particular, the symmetric group $S(n)$ acts simply-transitively on the set of Weyl chambers. This observation and the fact that $W \cong S(n)$ are no accident; we'll soon see that they are features of general compact connected Lie groups.

Let's briefly review the various actions at our disposal. We have actions of K on K by conjugation, on $\mathfrak{k}, \mathfrak{g}$ by the adjoint representation Ad, and on $\mathfrak{k}^*, \mathfrak{g}^*$ by the coadjoint representation Ad^*; we also have the differentiated actions ad of $\mathfrak{k}$ on $\mathfrak{k}, \mathfrak{g}$ (or of $\mathfrak{g}$ on $\mathfrak{g}$) and ad^* of $\mathfrak{k}$ on $\mathfrak{k}^*, \mathfrak{g}^*$ (or of $\mathfrak{g}$ on $\mathfrak{g}^*$). These are related as follows: if $g \in K$ and $x \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$, then

$$\langle x, \xi \rangle = \langle g \cdot x, g \cdot \xi \rangle;$$

here $\langle, \rangle$ denotes the canonical pairing $\mathfrak{g} \otimes \mathfrak{g}^* \to \mathbb{C}$ and we abbreviate $g \cdot x := \mathrm{Ad}(g)x$ and $g \cdot \xi := \mathrm{Ad}^*(g)\xi$. In terms of the differentiated actions, this reads: for $z \in \mathfrak{g}$, we have

$$\langle z \cdot x, \xi \rangle + \langle x, z \cdot \xi \rangle = 0,$$

where $\cdot$ denotes the evident action, e.g., $z \cdot x = \mathrm{ad}(z)x = [z, x]$.

The Weyl group $W = N(T)/T$ acts on T by conjugation, hence on $\mathfrak{t}, \mathfrak{t}_{\mathbb{R}}, \mathfrak{t}_{\mathbb{C}}$, etc., by the adjoint action inherited from G, and likewise on $\mathfrak{t}^*, \mathfrak{t}_{\mathbb{R}}^*$, etc., by the coadjoint action. The action of W on T induces an action of $w \in W$ on the cocharacter group $X^\vee(T)$ and on the character group $X(T)$, denoted $\gamma \mapsto w(\gamma)$ and $\lambda \mapsto w(\lambda)$, characterized in the first case by requiring that

$$z^{w(\gamma)} = w \cdot z^\gamma$$

(for $z \in \mathrm{U}(1)$ and $\gamma \in X^\vee(T)$) and in the second case by requiring that

$$(w \cdot t)^{w(\lambda)} = t^\lambda$$

(for $t \in T$ and $\lambda \in X(T)$). The actions of W on $X(T)$ and $X^\vee(T)$ and on $\mathfrak{t}_\mathbb{R}^*$ and $\mathfrak{t}_\mathbb{R}$ are compatible with the identifications $X(T) \cong \mathfrak{t}_\mathbb{Z}^* \subseteq \mathfrak{t}_\mathbb{R}^*$ and $X^\vee(T) \cong \mathfrak{t}_\mathbb{Z} \subseteq \mathfrak{t}_\mathbb{R}$, thus the actions of W on $\mathfrak{t}_\mathbb{R}$ and on $\mathfrak{t}_\mathbb{R}^*$ stabilize the lattices $\mathfrak{t}_\mathbb{Z}$ and $\mathfrak{t}_\mathbb{Z}^*$. Moreover:

Lemma 115. *(i) W acts on the set Φ of roots, i.e., $w(\alpha) \in \Phi$ for all $\alpha \in \Phi$; more precisely,*

$$w \cdot \mathfrak{g}^\alpha = \mathfrak{g}^{w(\alpha)}. \tag{51}$$

(ii) W acts on $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$.
(iii) W acts on the set of Weyl chambers, i.e., if C is a Weyl chamber, then so is $w(C)$ for each $w \in W$.

We might pause to convince ourselves that all of these properties are obvious in the familiar case of $\mathrm{U}(n)$.

Proof.   (i) We check (51). It suffices to show that $w \cdot \mathfrak{g}^\alpha$. (The same argument then gives $w^{-1} \cdot \mathfrak{g}^{w(\alpha)} \subseteq \mathfrak{g}^\alpha$, whence equality.) Let $x \in \mathfrak{g}^\alpha$ Then for $t \in T$,

$$t \cdot w \cdot x = \mathrm{Ad}(t)\,\mathrm{Ad}(w)x = \mathrm{Ad}(w)\,\mathrm{Ad}(w^{-1} \cdot t)x = (w^{-1} \cdot t)^\alpha\,\mathrm{Ad}(w)x = t^{w(\alpha)}\,\mathrm{Ad}(w)x.$$

Thus $\mathrm{Ad}(w)x \in \mathfrak{g}^{w(\alpha)}$, as required.

(ii) Recall that $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ consists of $x \in \mathfrak{t}_\mathbb{R}$ for which $\langle x, \alpha \rangle = \alpha(x)$ is nonzero for all roots α. We have $\langle w \cdot x, \alpha \rangle = \langle x, w^{-1}(\alpha) \rangle$, and we've seen that w^{-1} preserves the set of roots, so we conclude that w stabilizes $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$.

(iii) We use the previous assertion and the continuity of the action of W. $\qquad\square$

The following result illustrates well the power of knowing that tori have connected centralizers:

Lemma 116. *The action of W on the set of Weyl chambers is free, i.e., if C is a Weyl chamber and $w \in W$ satisfies $w(C) = C$, then $w = 1$.*

Proof. Let $W_C := \{w \in W : w(C) = C\}$. We must show that $W_C = \{1\}$. We've seen that C is a convex cone, and also that W and hence W_C is finite. Let $y_0 \in C$. Then the element $y := \sum_{w \in W_C} w \cdot y_0$ is a positive linear combination of elements of C, hence belongs to C. Let $w \in W_C$. Then w permutes the summands in the definition of y, so $w \cdot y = y$, i.e., $wT \subseteq Z_K(y)$. By Corollary 113, $Z_K(y)$ is connected. By general Lie theory, the Lie algebra of $Z_K(y)$ is given by $Z_\mathfrak{k}(y) = \{x \in \mathfrak{k} : [y, x] = 0\}$. Let $x \in Z_\mathfrak{k}(y)$. Writing $x = x_0 + \sum x_\alpha$ as in (49), we have $0 = [y, x] = 0 + \sum \alpha(y)x_\alpha$. Since $\alpha(y)$ is nonzero and the root space decomposition (43) is a direct sum decomposition, it follows that $x_\alpha = 0$ for all α, whence that $x = x_0 \in \mathfrak{t}_\mathbb{C} \cap \mathfrak{k} = \mathfrak{t}$. Therefore $Z_\mathfrak{k}(y) = \mathfrak{t}$. By the noted connectivity, it follows that $Z_K(y) = T$. Thus $wT \subseteq T$, i.e., $w = 1$ in W. We conclude as required that $W_C = \{1\}$. $\qquad\square$

We'll see later that the action of W on the set of Weyl chambers is moreover *transitive*; the above lemma will then tell us that it is simply-transitive, hence that the order $|W|$ of the Weyl group is equal to the number of Weyl chambers. So far, we haven't given any way to *produce* nontrivial elements of the Weyl group. (For instance, why should W be nontrivial when $K \neq T$?) That's our next main objective. The construction will involve the roots. For each root α, we'll cook up a map $F_\alpha : \mathrm{SU}(2) \to K$. We'll then use these maps to produce Weyl group elements, called *root reflections*, which we'll eventually show generate the Weyl group.

7.9. **Notation and preliminaries concerning** $\mathrm{SU}(2)$. Take $K = \mathrm{SU}(2)$, so that $\mathfrak{k} = \mathfrak{su}(2)$, $G = \mathrm{SL}_2(\mathbb{C})$, $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. Let $T \leq K$ denote the standard maximal torus $\{\mathrm{diag}(t, t^{-1}) : t \in \mathrm{U}(1)\}$, so that $\mathfrak{t} \leq \mathfrak{k}$ and $\mathfrak{t}_{\mathbb{C}} \leq \mathfrak{g}$ are the diagonal subalgebras. The map $\theta : \mathfrak{g} \to \mathfrak{g}$ given as usual by $\theta(x) = -x^* = -\overline{x}^t$ has the property that $\mathfrak{k}$ is the θ-fixed subspace of $\mathfrak{g}$. Define $X, Y, H \in \mathfrak{g}$ by

$$H := \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad X := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{52}$$

Then

$$[X, Y] = H, \quad [H, X] = 2X, \quad [H, Y] = -2Y \tag{53}$$

and

$$\theta(H) = -H, \quad \theta(X) = -Y, \quad \theta(Y) = -X. \tag{54}$$

These properties give a "presentation" for $\mathfrak{su}(2)$ in that for any pair $(\mathfrak{g}', \theta')$ consisting of a three-dimensional Lie algebra $\mathfrak{g}'$ equipped with an antilinear involution θ' that admits a basis H', X', Y' satisfying the analogue of (53) and (54), the map $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{g}'$ given by $H \mapsto H', X \mapsto X', Y \mapsto Y'$ is an isomorphism of complex Lie algebras that intertwines θ with θ', hence induces an isomorphism of real Lie algebras $\mathfrak{su}(2) \to \mathfrak{k}' := \{x \in \mathfrak{g}' : \theta'(x) = x\}$.

We have

$$\mathfrak{t}_{\mathbb{C}} = \mathbb{C}H, \quad \mathfrak{t}_{\mathbb{R}} = \mathbb{R}H, \quad \mathfrak{t}_{\mathbb{Z}} = \mathbb{Z}H.$$

(For the last of these, we note that every one-parameter subgroup $\mathrm{U}(1) \to T$ is given by $z \mapsto \mathrm{diag}(z^m, z^{-m})$ for some $m \in \mathbb{Z}$; we may rewrite this as $\exp(i\theta) \mapsto \exp(imH)$, whence the claim.) In view of (53), the root space decomposition of $\mathfrak{g}$ is given by

$$\mathfrak{g} = \underbrace{\mathbb{C}H}_{\mathfrak{t}_{\mathbb{C}}} \oplus \underbrace{\mathbb{C}X}_{\mathfrak{g}^{\alpha}} \oplus \underbrace{\mathbb{C}Y}_{\mathfrak{g}^{-\alpha}},$$

where $\alpha \in \mathfrak{t}_{\mathbb{R}}^*$ is characterized by

$$\alpha(H) = 2.$$

The set of roots is thus

$$\Phi = \{\alpha, -\alpha\}.$$

We note that

$$\mathfrak{t}_{\mathbb{C}}^* = \mathbb{C}\alpha, \quad \mathfrak{t}_{\mathbb{R}}^* = \mathbb{R}\alpha, \quad \mathfrak{t}_{\mathbb{Z}}^* = \tfrac{1}{2}\mathbb{Z}\alpha.$$

(For the last of these, note that the characters of T are given by $\mathrm{diag}(t, t^{-1}) \mapsto t^l$ for some $l \in \mathbb{Z}$, and that if we write $\mathrm{diag}(t, t^{-1}) = \exp(i\theta H)$, then $t^l = \exp(i\theta l) = \exp(i\theta l\alpha(H)/2) = \exp(i\theta H)^{l\alpha/2}$, whence the characters of T are uniquely of the form $l\alpha/2$ with $l \in \mathbb{Z}$.)

Recall (§3.8 and earlier) that any finite-dimensional representation $\pi : \mathrm{SU}(2) \to \mathrm{GL}(V)$ decomposes as a sum

$$V = \oplus_{k \in \mathbb{Z}} V^k$$

of weight spaces

$$V^k := \{v \in V : \mathrm{diag}(t, t^{-1})v = t^k v \text{ for all } t \in \mathrm{U}(1)\}.$$

For example, if $(\pi, V) = (\mathrm{Ad}, \mathfrak{g})$, then $V^0 = \mathbb{C}H = \mathfrak{t}_{\mathbb{C}}, V^2 = \mathbb{C}X = \mathfrak{g}^{\alpha}, V^{-2} = \mathbb{C}Y = \mathfrak{g}^{-\alpha}$. In terms of the differentiated representation $d\pi : \mathfrak{su}(2) \to \mathrm{End}(V)$ and its complexification $d\pi : \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{End}(V)$, we may write

$$V^k = \{v \in V : d\pi(H)v = kv\},$$

that is to say, the weight space decomposition of (π, V) is simply its decomposition into eigenspaces for the operator $d\pi(H)$. Note in particular that every eigenvalue of $d\pi(H)$ is an *integer*.

For simplicity of notation we will often use abbreviations like $xv := d\pi(x)v$ for $x \in \mathfrak{g}$ and $v \in V$ when π is clear from context.

In view of the general identity "$\mathfrak{g}^\alpha \cdot V^\lambda \subseteq V^{\lambda+\alpha}$" (Lemma 92) and the fact that $\alpha(H) = 2$, we see that X and Y act on the H-eigenspaces by "raising and lowering operators," i.e.,

$$d\pi(X) : V^k \to V^{k+2}, \quad d\pi(Y) : V^k \to V^{k-2}.$$

For instance, this follows in the case $(\pi, V) = (\mathrm{Ad}, \mathfrak{g})$ from the identities (53).

We have also seen that the irreducible finite-dimensional representation of $\mathrm{SU}(2)$ are indexed by nonnegative integers n and given explicitly by the $(n{+}1)$-dimensional space $V_n = \mathbb{C}[x, y]^{(n)}$ of homogeneous polynomials of degree n in two variables, with $\mathrm{SU}(2)$ acting by right translation. Moreover, any finite-dimensional representation (π, V) of $\mathrm{SU}(2)$ decomposes uniquely as a finite direct sum $V = \oplus_{n \geq 0} V_n^{\oplus \mu(n)}$ for some multiplicities $\mu(n) \geq 0$,

The weight spaces of the irreducible representations (π_n, V_n) are all one-dimensional, and the weight space decompositions read

$$V_n = V_n^n \oplus V_n^{n-2} \oplus V_n^{n-2} \oplus \cdots \oplus V_n^{-n}$$
$$= \mathbb{C}x^n \oplus \mathbb{C}x^{n-1}y \oplus \mathbb{C}x^{n-2}y^2 \oplus \cdots \oplus \mathbb{C}y^n.$$

Using the definition (52), we compute that the action of $\mathfrak{g}$ is given by

$$d\pi_n(H) = x\partial_x - y\partial_y, \quad d\pi_n(X) = x\partial_y, \quad d\pi_n(Y) = y\partial_x.$$

It follows readily from this explicit description that the lowering maps

$$V_n^n \xrightarrow{Y} V_n^{n-2} \xrightarrow{Y} V_n^{n-2} \xrightarrow{Y} \cdots \xrightarrow{Y} V_n^{-n}$$

and the raising maps

$$V_n^{-n} \xrightarrow{X} V_n^{-n+2} \xrightarrow{X} V_n^{-n+4} \xrightarrow{X} \cdots \xrightarrow{X} V_n^n \tag{55}$$

are all isomorphisms of one-dimensional vector spaces. Indeed, they send each monomial to an explicit nonzero multiple of another monomial. Of course $d\pi(X)V_n^n$ and $d\pi(Y)V_n^{-n}$ are trivial, because the weight spaces V_n^{n+2} and V_n^{-n-2} are trivial.

Here's a typical application of the above explication of the representation theory of $\mathrm{SU}(2)$; this application will turn out to be the key for showing that the root spaces of a general compact Lie group are one-dimensional.

Lemma 117. *Let V be a finite-dimensional representation of $\mathrm{SU}(2)$ such that*

 (i) *all weights (i.e., eigenvalues of H) are even, and*
(ii) *the weight space V^0 is one-dimensional.*

Then V is irreducible, say $V \cong V_n$. If moreover there exists a nonzero vector $v \in V^2$ such that $Xv = 0$, then $n = 2$.

Proof. Since the weights are even, when we decompose $V = \oplus V_n^{\oplus \mu(n)}$ as a sum of irreducibles V_n, only those with n even occur with positive multiplicity $\mu(n)$. For n even, we have $\dim V_n^0 = 1$, thus $1 = \dim V^0 = \sum_n \mu(n)$. Thus $\mu(n) = 1$ for some n and vanishes for all other values, and so $V \cong V_n$.

Suppose there exists v with the stated property. Since v is nonzero and X annihilates v, we see (from the assertion following (55)) that $n = 2$. □

[14]

Similarly:

Lemma 118. *Let V be a finite-dimensional representation of $\mathrm{SU}(2)$ that has some odd weight, i.e., the weight space V^n is nonzero for some odd integer n. Then the weight space V^1 is nonzero.*

Proof. We decompose $V = \oplus_{n \geq 0} V_n^{\oplus \mu(n)}$ into irreducibles. Since the weights of V_n all have the same parity as n, we see that $\mu(n)$ is positive for some odd n. But for such n, the weights of V_n are $n, n-2, n-4, \ldots, 3, 1, -1, -3, \ldots, -n$. Thus $\dim V_n^1 > 0$ and so $\dim V^1 \geq \mu(n) \dim V_n^1 > 0$, as required. □

7.10. **From roots to distinguished** $\mathrm{SU}(2)$**'s.** We're now ready to state and prove one of the most important theorems concerning the structure of compact Lie groups.

Recall the notation (52) for the standard basis elements H, X, Y of $\mathfrak{sl}_2(\mathbb{C})$.

Theorem 119. *Let K be a compact connected Lie group, $T \leq K$ a maximal torus, $\Phi = \Phi(K : T)$ the associated set of roots, and $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$ the complexified Lie algebra with the usual root space decomposition. Let $\alpha \in \Phi \subseteq \mathfrak{t}_{\mathbb{Z}}^*$ be a root.*

(1) There is a morphism of Lie groups

$$F_\alpha : \mathrm{SU}(2) \to K$$

whose differential $dF_\alpha : \mathfrak{su}(2) \to \mathfrak{k}$ has complexification $dF_\alpha : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{g}$ satisfying

$$dF_\alpha(X) \in \mathfrak{g}^\alpha.$$

(2) There is a cocharacter $\alpha^\vee \in X^\vee(T) = \mathrm{Hom}(\mathrm{U}(1), T)$ so that for each such F_α, we have a commutative diagram

$$\begin{array}{ccc}
\mathrm{U}(1) & \xrightarrow{\alpha^\vee} & T \\
{\scriptstyle t \mapsto \mathrm{diag}(t, 1/t)} \downarrow & & \downarrow \\
\mathrm{SU}(2) & \xrightarrow{F_\alpha} & K.
\end{array} \qquad (56)$$

The cocharacter $\alpha^\vee$ is independent of the choice of F_α. More precisely, let

$$H_\alpha \in \mathfrak{t}_{\mathbb{Z}} \cong X^\vee(T)$$

correspond to $\alpha^\vee$ in the sense of §7.1, thus

$$\alpha^\vee(e^{i\theta}) = \exp(i\theta H_\alpha)$$

for all $\theta \in \mathbb{R}$. Then H_α is the unique element of $[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$ such that $\alpha(H_\alpha) = 2$.

(3) $\dim(\mathfrak{g}^\alpha) = 1$.
(4) $\mathbb{C}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi = \{\alpha, -\alpha\}$.
(5) F_α is unique up to conjugation by the image of $\alpha^\vee$.

---

[14]End of lecture #13, Thursday, 4 Apr

(6) Let $T_\alpha \leq T$ denote the codimension one subtorus defined as the connected component of the kernel of α, thus

$$T_\alpha := \ker(\alpha)^0 = \{t \in T : t^\alpha = 1\}^0, \quad \mathfrak{t}_\alpha := \mathrm{Lie}(T_\alpha) = \{x \in \mathfrak{t} : \alpha(x) = 0\}.$$

Then the image of F_α is contained in the centralizer $Z_K(T_\alpha)$, and the composition

$$\mathrm{SU}(2) \xrightarrow{F_\alpha} Z_K(T_\alpha) \to Z_K(T_\alpha)/T_\alpha$$

is surjective, and induces an isomorphism $\mathfrak{su}(2) \cong Z_{\mathfrak{k}}(\mathfrak{t}_\alpha)/\mathfrak{t}_\alpha$ of Lie algebras.

The point is that a root α on its own is fairly useless – there's not much you can do – but once you know that it comes from some copy of $\mathrm{SU}(2)$, everything becomes possible. If we learn anything from this part of the course, it's that *we should always think of a root of a compact connected Lie group as coming with an associated copy of* $\mathrm{SU}(2)$ in the above sense.

Let's illustrate with the example of $K = \mathrm{U}(3)$, with $T = \mathrm{U}(1)^3$ the standard diagonal maximal torus and $\alpha = \varepsilon_1 - \varepsilon_2$, so that $\mathfrak{g}^\alpha = \mathbb{C}E_{12}$. We may then take

$$F_\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so that $dF_\alpha(X) = E_{12} \in \mathfrak{g}^\alpha$. The composition of the standard cocharacter $\mathrm{U}(1) \xrightarrow{t \mapsto \mathrm{diag}(t,1/t)} \mathrm{SU}(2)$ with F_α is the cocharacter

$$\alpha^\vee : \mathrm{U}(1) \to T$$

$$t \mapsto \begin{pmatrix} t & & \\ & 1/t & \\ & & 1 \end{pmatrix}$$

of K, and we have $\alpha^\vee(e^{i\theta}) = \exp(i\theta H_\alpha)$ with

$$H_\alpha = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} = dF_\alpha(H).$$

We have $\Phi = \{\varepsilon_i - \varepsilon_j : i \neq j\}$, so it is clear that $\mathbb{Q}\alpha \cap \Phi = \{\alpha, -\alpha\}$. We compute using the Lie algebra that

$$T_\alpha = \left\{ \begin{pmatrix} z & & \\ & z & \\ & & t \end{pmatrix} : z, t \in \mathrm{U}(1) \right\},$$

$$Z_K(T_\alpha) = \left\{ \begin{pmatrix} a & b & \\ c & d & \\ & & t \end{pmatrix} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{U}(2), t \in \mathrm{U}(1) \right\}.$$

We may verify that $F_\alpha : \mathrm{SU}(2) \to Z_K(T_\alpha)/T_\alpha$ is surjective with kernel $\{\pm 1\}$, and in any event that dF_α defines an isomorphism of Lie algebras (which is simplest to check first for the complexified Lie algebras).

If we were instead given $\alpha = \varepsilon_1 - \varepsilon_3$, then we'd take

$$F_\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ c & 0 & d \end{pmatrix}.$$

7.11. **Classification of rank one groups.** Before proving Theorem 119, let's illustrate its power with a typical application.

Definition 120. The *rank* of a compact connected Lie group K is defined to be the dimension of any maximal torus T (which is well-defined by Theorem 107, for instance):

$$\mathrm{rank}(K) := \dim(T).$$

For instance, $\mathrm{rank}(\mathrm{U}(n)) = n$, while $\mathrm{rank}(\mathrm{SU}(n)) = n - 1$.

It's clear from Lemma 89 that every connected compact Lie group of rank 0 is trivial. The case of rank 1 is more interesting:

Corollary 121. *Every connected compact Lie groups K of rank 1 is isomorphic to exactly one of the following:*

- $\mathrm{U}(1)$
- $\mathrm{SU}(2)$
- $\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}(3)$

Proof. Let $T \leq K$ be a maximal torus and $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$, with Φ the set of roots. If Φ is empty, then the root space decomposition implies that $\mathfrak{g} = \mathfrak{t}_{\mathbb{C}}$, hence that $\mathfrak{k} = \mathfrak{t}$; since K is connected, this forces $K = T \cong \mathrm{U}(1)$. Suppose otherwise that there exists some $\alpha \in \Phi$. Then the codimension 1 subtorus $T_{\alpha} \leq T$ defined in Theorem 119 has dimension $\dim(T_{\alpha}) = \dim(T) - 1 = 0$, and so $T_{\alpha} = \{1\}$. Thus $Z_K(T_{\alpha}) = K$ and $Z_K(T_{\alpha})/T_{\alpha} \cong K$. Thus any map $F_{\alpha} : \mathrm{SU}(2) \to K$ as in Theorem 119 defines a Lie algebra isomorphism $\mathfrak{su}(2) \to K$. Since K is connected, F_{α} is surjective. Since dF_{α} is injective, the kernel of F_{α} is a (discrete) subgroup of the center $\{\pm 1\}$ of $\mathrm{SU}(2)$ (see e.g. §20 and especially Lemma 154 of my notes on Lie groups, on the course homepage; the general fact being used here is that if $f : G \to H$ is a morphism of Lie groups with df injective and G connected, then $\ker(f)$ is a discrete subgroup of the center of G). Thus either $K \cong \mathrm{SU}(2)$ or $K \cong \mathrm{SU}(2)/\{\pm 1\}$; the two cases are distinguished by (for instance) their centers. $\square$

In fact, this application is representative of the power of Theorem 119 in the sense that one could go backwards and derive much of Theorem 121 assuming Corollary 121. We will not do this here, but it's done that way in some references, e.g., the BTD course reference.

7.12. **Some invariant inner products.** Before turning to the proof of Theorem 119, we pause to choose some inner products concerning the objects in its statement. (The formulation of Theorem 119 is independent of this choice, but the proof will not be.) Choose an embedding $K \hookrightarrow \mathrm{U}(n)$ with respect to which T is contained in the diagonal subgroup (see §7.1 for details on this and what follows). This choice defines an embedding $\mathfrak{g} \hookrightarrow \mathfrak{gl}_n(\mathbb{C})$ with respect to which $\mathfrak{t}_{\mathbb{R}}$ is contained in the subspace of diagonal matricse with real entries. We have the usual involution $\theta : \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$, preserving $\mathfrak{g}$, and given by $\theta(x) := -x^*$, where $x^* := \overline{x}^t$. Then:

- The map

$$\mathfrak{gl}_n(\mathbb{C}) \times \mathfrak{gl}_n(\mathbb{C}) \to \mathbb{R}$$

$$(x, y) \mapsto -\mathrm{trace}(x\theta(y))$$

defines an inner product. Indeed,

$$- \operatorname{trace}(x\theta(y)) = \operatorname{trace}(xy^*) = \sum_{i,j} x_{ij}\overline{y_{ij}}.$$

In particular, this map restricts to an inner product on $\mathfrak{g}$.

- The map

$$\mathfrak{t}_{\mathbb{R}} \times \mathfrak{t}_{\mathbb{R}} \to \mathbb{R}$$

$$(x, y) \mapsto \operatorname{trace}(xy)$$

is an inner product. Indeed, for $y \in \mathfrak{t}_{\mathbb{R}}$, we have $y = -\theta(y)$, so this is just the restriction of the inner product defined above on $\mathfrak{gl}_n(\mathbb{C})$; explicitly, it is just the obvious inner product given in terms of the diagonal entries by the formula $\operatorname{trace}(xy) = \sum_i x_{ii}\overline{y_{ii}}$.

- The inner product just defined on $\mathfrak{t}_{\mathbb{R}}$ defines an isomorphism

$$\mathfrak{t}_{\mathbb{R}}^* \to \mathfrak{t}_{\mathbb{R}}$$

$$\lambda \mapsto u_\lambda$$

characterized by the identity

$$\lambda(z) = \operatorname{trace}(u_\lambda z) \tag{57}$$

for all $z \in \mathfrak{t}_{\mathbb{R}}$. The same identity extends $\mathbb{C}$-linearly to all $z \in \mathfrak{t}_{\mathbb{C}}$.

- Using the above isomorphism, we may transport the given inner product on $\mathfrak{t}_{\mathbb{R}}$ to an inner product on $\mathfrak{t}_{\mathbb{R}}^*$ that we denote by $(,)$: for $\lambda_1, \lambda_2 \in \mathfrak{t}_{\mathbb{R}}^*$,

$$(\lambda_1, \lambda_2) := \operatorname{trace}(u_{\lambda_1} u_{\lambda_2}).$$

For example, if $K = \mathrm{U}(n)$, then under the identification $\mathfrak{t}_{\mathbb{R}}^* \cong \mathbb{R}^n$ as in §7.1, this inner product is the standard one.

We verify readily that the inner product that we defined on $\mathfrak{gl}_n(\mathbb{C})$ is $\mathrm{Ad}(\mathrm{U}(n))$-invariant, hence that on $\mathfrak{g}$ is $\mathrm{Ad}(K)$-invariant and that on $\mathfrak{t}_{\mathbb{R}}$ is W-invariant.

7.13. **Proofs concerning the passage from roots to** $\mathrm{SU}(2)$**'s.** With these preliminaries out of the way, we're now prepared to give the:

Proof of Theorem 119.   (i) We'll construct F_α in terms of generators and relations (see (53) and (54)). We start by constructing elements $X_\alpha, Y_\alpha, H_\alpha \in \mathfrak{g}$ that we eventually intend to be the images of $X, Y, H \in \mathfrak{sl}_2(\mathbb{C})$ under dF_α.

First, let $c > 0$ denote a suitable constant to be specified later (in (65)), and choose any $X_\alpha \in \mathfrak{g}^\alpha$ whose norm $\|X_\alpha\|^2 := -\operatorname{trace}(X_\alpha\theta(X_\alpha))$ is equal to c. Set $Y_\alpha := -\theta(X_\alpha) \in \mathfrak{g}^{-\alpha}$ and $H_\alpha := [X_\alpha, Y_\alpha] \in \mathfrak{t}_{\mathbb{C}}$. We get a linear map

$$\text{``} dF_\alpha'' : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{g}$$

$$X \mapsto X_\alpha, Y \mapsto Y_\alpha, H \mapsto H_\alpha.$$

We want the above map to define a θ-equivariant Lie algebra morphism. For this to be the case, the non-obvious relations to be verified are that

$$[H_\alpha, X_\alpha] = 2X_\alpha, \tag{58}$$

$$[H_\alpha, Y_\alpha] = -2Y_\alpha, \tag{59}$$

$$\theta(H_\alpha) = -H_\alpha. \tag{60}$$

Since $[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha$ and $[H_\alpha, Y_\alpha] = -\alpha(H_\alpha)Y_\alpha$, the two relations (59) are equivalent to the single relation

$$\alpha(H_\alpha) = 2, \tag{61}$$

while the third relation (60) is equivalent to requiring that

$$H_\alpha \in \mathfrak{t}_\mathbb{R}. \tag{62}$$

Before checking the above relations, we note that for all $x \in \mathfrak{g}^\alpha, y \in \mathfrak{g}^{-\alpha}$ and $z \in \mathfrak{t}_\mathbb{C}$, we have

$$\operatorname{trace}([x, y]z) = \alpha(z) \operatorname{trace}(xy). \tag{63}$$

Indeed, since the trace of a product of two matrices doesn't change if we swap the two matrices, we have $\operatorname{trace}([x, y]z) = \operatorname{trace}(z[x, y]) = \operatorname{trace}(zxy - zyx) = \operatorname{trace}(zxy - xzy) = \operatorname{trace}([z, x]y)$, and since $[z, x] = \alpha(z)$ the required identity follows. By the characterizing property (57) of u_α, we may rewrite (63) as the identity

$$[x, y] = \operatorname{trace}(xy)u_\alpha \tag{64}$$

of elements of $\mathfrak{t}_\mathbb{C}$.

Applying (64) to X_α and Y_α gives

$$H_\alpha = [X_\alpha, Y_\alpha] = \operatorname{trace}(X_\alpha Y_\alpha)u_\alpha = \|X_\alpha\|^2 u_\alpha = cu_\alpha.$$

using in the last two steps the definitions of $\|X_\alpha\|^2$ and of Y_α. On the other hand, the definition of u_α gives

$$\alpha(u_\alpha) = \operatorname{trace}(u_\alpha u_\alpha) = (\alpha, \alpha).$$

Combining the above two identites, we obtain $\alpha(H_\alpha) = c(\alpha, \alpha)$. With the choice

$$c := \frac{2}{(\alpha, \alpha)}, \tag{65}$$

we obtain the required relation (61). Moreover, we have $H_\alpha = cu_\alpha \in \mathbb{R}u_\alpha \subseteq \mathfrak{t}_\mathbb{R}$, so (62) holds.

This completes the proof that the linear map "dF_α" defines a θ-equivariant $\mathbb{C}$-linear Lie algebra morphism $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{g}$. By passage to θ-fixed subspaces, we obtain a Lie algebra morphism $\mathfrak{su}(2) \to \mathfrak{k}$. Since SU(2) is simply-connected, this lifts to a Lie group morphism $F_\alpha : K \to \mathrm{SU}(2)$, whose differential dF_α is the map "dF_α" that we constructed.

(ii) The composition

$$\alpha^\vee : \mathrm{U}(1) \to \mathrm{SU}(2) \xrightarrow{F_\alpha} K$$

is given by

$$e^{i\theta} \mapsto \exp(i\theta dF_\alpha(H)) = \exp(i\theta H_\alpha) \in T,$$

so the diagram (56) is defined and commutes. Finally, we see from (64) that

$$[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}] \subseteq \mathbb{C}u_\alpha = \mathbb{C}H_\alpha, \tag{66}$$

so that $H_\alpha = [X_\alpha, Y_\alpha]$ admits the required characterization as the unique element of $[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$ for which $\alpha(H_\alpha) = 2$.

(iii) Define

$$V := \mathbb{C}H_\alpha \oplus (\oplus_{0 \neq n \in \mathbb{Z}}\mathfrak{g}^{n\alpha}), \tag{67}$$

and set $\mathfrak{sl}_2(\mathbb{C})_\alpha := dF_\alpha(\mathfrak{sl}_2(\mathbb{C}))$. We may verify then $\mathfrak{sl}_2(\mathbb{C})_\alpha$ that acts $\mathbb{C}$-linearly on V by the restriction of the adjoint representation ad:

- $\operatorname{ad}_{H_\alpha}$ preserves each summand in (67), acting on $\mathbb{C}H_\alpha$ by the eigenvalue 0 and on each $\mathfrak{g}^{n\alpha}$ by the eigenvalue $n\alpha(H_\alpha) = 2n$.
- $\operatorname{ad}_{X_\alpha}(\mathfrak{g}^{n\alpha}) \subseteq \mathfrak{g}^{(n+1)\alpha}$ for all integers n. Moreover, $\operatorname{ad}_{X_\alpha}(\mathfrak{g}^{-\alpha}) \subseteq \mathbb{C}H_\alpha$ thanks to (66). Thus $\operatorname{ad}_{X_\alpha}$ maps V to V.

- Similarly, $\mathrm{ad}_{Y_\alpha}(\mathfrak{g}^{n\alpha}) \subseteq \mathfrak{g}^{(n-1)\alpha}$ for all integers n, and $\mathrm{ad}_{Y_\alpha}(\mathfrak{g}^\alpha) \subseteq \mathbb{C}H_\alpha$. Thus V defines a finite-dimensional $\mathbb{C}$-linear representation of $\mathfrak{sl}_2(\mathbb{C})$, hence by Theorem (86) a finite-dimensional representation of SU(2). This representation has the property that all weights (i.e., H_α-eigenvalues) are even, the weight zero subspace $\mathbb{C}H_\alpha$ is one-dimensional, and the weight two subspace $\mathfrak{g}^\alpha$ contains the vector X_α for which $\mathrm{ad}_{X_\alpha} X_\alpha = 0$. By Lemma 117, we deduce that V is isomorphic to the irreducible three-dimensional representation V_2 of SU(2), whose weights are $\{0, 2, -2\}$ each occurring with multiplicity one. Thus

$$\dim \mathfrak{g}^\alpha = \dim \mathfrak{g}^{-\alpha} = 1$$

and

$$\dim \mathfrak{g}^{n\alpha} = 0 \text{ for } n \in \mathbb{Z} - \{0, 1, -1\}. \tag{68}$$

(iv) We now take instead

$$V = \mathbb{C}H_\alpha \oplus (\oplus_{0 \neq n \in \mathbb{C}\alpha} \mathfrak{g}^{n\alpha}).$$

(We might note in passing that, since the set Φ of roots is contained in the finite free $\mathbb{Z}$-module $\mathfrak{t}_\mathbb{Z}^*$, the sum over n could be restricted from the start to $\mathbb{Q}\alpha$.) We note as before that $\mathfrak{sl}_2(\mathbb{C})_\alpha$ acts on V by the restriction of ad. Then V defines a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$, and so, as we have seen, all of the weights of ad_{H_α} must be integers. But the eigenvalue of ad_{H_α} on $\mathfrak{g}^{n\alpha}$ is $n\alpha(H_\alpha) = 2n$, which is an integer only if $n \in (1/2)\mathbb{Z}$. We've seen already in (68) that for $n \in \mathbb{Z}$, we have $n\alpha \notin \Phi$ unless $n = \pm 1$, so it remains only to consider the case that $n\alpha \in \Phi$ for some $n \in (1/2)\mathbb{Z} - \mathbb{Z}$. But then $2n$ is odd, so $\mathfrak{g}^{n\alpha}$ is a nonzero weight space for V with odd weight, so by Lemma 118, we see that the weight one subspace $V^1 = \mathfrak{g}^{\alpha/2}$ is nonzero, i.e., that $\alpha/2$ is a root. But then $\alpha/2$ and $\alpha = 2(\alpha/2)$ are both roots, which contradicts (68) (applied to $\alpha/2$ in place of α). The proof that $\mathbb{C}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi = \{\alpha, -\alpha\}$ is now complete.[15]

(v) The uniqueness of F_α up to conjugation by $\alpha^\vee$ follows from the recently established one-dimensionality of $\mathfrak{g}^\alpha$ and the construction of F_α. Since we required $\|X_\alpha\|^2 = 2/(\alpha, \alpha)$ and the choices of Y_α and H_α were then forced by the requirement that dF_α be a θ-equivariant Lie algebra morphism, any other map F'_α satisfying the same conditions as F_α is obtained by replacing X_α as in the construction of F_α with its multiple by some element $e^{i\theta} \in \mathrm{U}(1)$. Since $[H_\alpha, X_\alpha] = 2X_\alpha$, we have

$$\mathrm{Ad}(\alpha^\vee(e^{i\theta/2}))X_\alpha = \mathrm{Ad}(\exp(i\theta H_\alpha/2))X_\alpha = e^{i\theta}X_\alpha,$$

which implies that F'_α is the conjugate of F_α by $\alpha^\vee(e^{i\theta/2})$, as required.

(vi) We have $\mathfrak{t}_\alpha = \{x \in \mathfrak{t} : \alpha(x) = 0\}$ and and $\mathrm{Lie}(Z_K(T_\alpha)) = Z_\mathfrak{k}(\mathfrak{t}_\alpha)$ and $(Z_\mathfrak{k}(\mathfrak{t}_\alpha))_\mathbb{C} = Z_\mathfrak{g}(\mathfrak{t}_\alpha) = \{x \in \mathfrak{g} : [x, z] = 0 \text{ for all } z \in \mathfrak{t} \text{ with } \alpha(z) = 0\}$. Since $[z, X_\alpha] = \alpha(z)X_\alpha$ and $[z, Y_\alpha] = -\alpha(z)Y_\alpha$ for all $z \in \mathfrak{t}$, and since $\mathfrak{t}_\mathbb{C}$ is commutative, we see that $X_\alpha, Y_\alpha, H_\alpha$ all belong to $Z_\mathfrak{g}(\mathfrak{t}_\alpha)$. Since SU(2) is connected, it follows likewise that the image of F_α is contained in $Z_K(T_\alpha)$. It remains only to check that the map

$$dF_\alpha : \mathfrak{su}(2) \to Z_\mathfrak{k}(\mathfrak{t}_\alpha)/\mathfrak{t}_\alpha$$

---

[15]End of lecture #14, Tuesday, 9 Apr

is an isomorphism, or equivalently, that

$$dF_\alpha : \mathfrak{sl}_2(\mathbb{C}) \to Z_\mathfrak{g}(\mathfrak{t}_\alpha)/(\mathfrak{t}_\alpha)_\mathbb{C} \tag{69}$$

is an isomorphism. We note that $Z_\mathfrak{g}(\mathfrak{t}_\alpha)$ is a Lie subalgebra of $\mathfrak{g}$ that contains $\mathfrak{t}_\mathbb{C}$, hence admits a root space decomposition

$$Z_\mathfrak{g}(\mathfrak{t}_\alpha) = \mathfrak{t}_\mathbb{C} \oplus (\oplus_{\beta \in B} \mathfrak{g}^\beta),$$

where B denotes the set of all roots $\beta \in \Phi$ for which $\mathfrak{g}^\beta \subseteq Z_\mathfrak{g}(\mathfrak{t}_\alpha)$, i.e., for which $\beta(z) = 0$ for all $z \in \mathfrak{t}$ with $\alpha(z) = 0$. Since α and β are both nonzero functionals, it follows that $\beta \in B$ iff $\beta \in \mathbb{C}\alpha$, which we've seen is equivalent to asking that $\beta = \pm\alpha$. Since $\alpha(H_\alpha) \neq 0$, we have a decomposition

$$\mathfrak{t}_\mathbb{C} = (\mathfrak{t}_\alpha)_\mathbb{C} \oplus \mathbb{C}H_\alpha. \tag{70}$$

Thus

$$Z_\mathfrak{g}(t_\alpha) = (\mathfrak{t}_\alpha)_\mathbb{C} \oplus \mathbb{C}H_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}$$

and so

$$Z_\mathfrak{g}(\mathfrak{t}_\alpha)/(\mathfrak{t}_\alpha)_\mathbb{C} \cong \mathbb{C}H_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}.$$

The composition

$$\mathfrak{sl}_2(\mathbb{C}) \xrightarrow{dF_\alpha} Z_\mathfrak{g}(t_\alpha) \twoheadrightarrow Z_\mathfrak{g}(t_\alpha)/(t_\alpha)_\mathbb{C} \cong \mathbb{C}H_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha} \cong \mathfrak{sl}_2(\mathbb{C})$$

is the identity, so we deduce as required that (69) is an isomorphism.

$\square$

7.14. **Construction of root reflections.** We finally develop a way to produce nontrivial elements of the Weyl group. Start with $K = \mathrm{SU}(2)$, with $T \cong \mathrm{U}(1)$ the standard diagonal maximal torus. It follows from our earlier analysis of $\mathrm{U}(n)$ that the Weyl group W of T is isomorphic to $\mathbb{Z}/2$. One can see this by noting that T consists of those elements of K that are diagonalized by the decomposition $\mathbb{C}^2 = \mathbb{C}e_1 \oplus \mathbb{C}e_2$ of the standard representation into the lines spanned by its standard basis elements. Hence $N(T)$ acts faithfully on such lines, with T acting trivially. Thus W embeds as a subgroup of the permutation group of $\{\mathbb{C}e_1, \mathbb{C}e_2\}$. The element $w := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \in K$ swaps the two lines, so in fact $W = \{eT, wT\}$ is the full permutation group. The same proof shows that the Weyl group of each of $\mathrm{U}(n), \mathrm{SU}(n), (n)$ is the symmetric group $S(n)$. Anyway, we compute that with $H = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{C})$,

$$w \cdot H = -H.$$

More generally, w acts on each of the spaces $\mathfrak{t}, \mathfrak{t}_\mathbb{Z}, \mathfrak{t}_\mathbb{R}, \mathfrak{t}_\mathbb{C}, \mathfrak{t}^*, \mathfrak{t}^*_\mathbb{Z}, \mathfrak{t}^*_\mathbb{R}, \mathfrak{t}^*_\mathbb{C}$ by negation. Moreover, for any finite-dimensional representation V of $\mathrm{SU}(2)$ with weight space decomposition $V = \oplus_{\lambda \in \mathbb{Z}} V^\lambda$ (here $V^\lambda = \{v \in V : Hv = \lambda v\}$, as usual), the action of w defines an isomorphism

$$w : V^\lambda \xrightarrow{\cong} V^{-\lambda},$$

as follows from a routine computation as in the proof of Lemma 115.

Returning now to the general setting:

Theorem 122. *Let K be a compact connected Lie group, with maximal torus T. Let $\Phi = \Phi(K : T)$ denote the set of roots, and $W = N(T)/T$ the Weyl group.*

Let $\alpha \in \Phi$. Let $F_\alpha : \mathrm{SU}(2) \to K$, as in Theorem 119, be such that $F_\alpha(\begin{pmatrix} & 1 \\ & \end{pmatrix})$ is a nonzero element of $\mathfrak{g}^\alpha$. Define

$$w_\alpha := F_\alpha(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}).$$

Then:

(i) *w_α belongs to the normalizer $N(T)$ of T, hence defines a Weyl group element that we denote by*
$$s_\alpha := w_\alpha T \in W.$$
The element s_α is independent of the choice of F_α.

(ii) *$s_\alpha^2 = 1$.*

(iii) *For $x \in \mathfrak{t}_\mathbb{C}$,*
$$s_\alpha \cdot x = x - \alpha(x)H_\alpha$$
and for $t \in T$,
$$s_\alpha \cdot t = t/\alpha^\vee(t^\alpha).$$

(iv) *For $\lambda \in \mathfrak{t}_\mathbb{C}^*$,*
$$s_\alpha \cdot \lambda = \lambda - \lambda(H_\alpha)\alpha.$$

Proof. We check readily that assertions (iii) and (iv) are equivalent, that either implies (ii), and that the claimed uniqueness of s_α follows from the established uniqueness of F_α. On the other hand, each of the assertions (i) and (iii) will follow if we can show that
$$\mathrm{Ad}(w_\alpha)x = x - \alpha(x)H_\alpha$$
for all $x \in \mathfrak{t}_\mathbb{C}$. To that end, we apply the decomposition (70) to write $x = y + zH_\alpha$, where $y \in \mathfrak{t}_\mathbb{C}$ with $\alpha(y) = 0$ and $z \in \mathbb{C}$, namely $z = (1/2)\alpha(x)$. By part (vi) of Theorem 119, the image of F_α commutes with y, hence $\mathrm{Ad}(w_\alpha)y = y$. Set $w := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \in \mathrm{SU}(2)$. Then

$$\mathrm{Ad}(w_\alpha)H_\alpha = \mathrm{Ad}(F_\alpha(w))F_\alpha(H) = F_\alpha(\mathrm{Ad}(w)H) = F_\alpha(-H) = -H_\alpha,$$

so

$$\mathrm{Ad}(w_\alpha)x = y - zH_\alpha = x - 2zH_\alpha = x - \alpha(x)H_\alpha,$$

as required. $\qquad\square$

[16]

The theorem and its proof imply that the action of s_α on $\mathfrak{t}_\mathbb{R}$ may be characterized as the unique linear map that restricts to the identity on $\ker(\alpha)$ and sends H_α to $-H_\alpha$; similarly, the action on $\mathfrak{t}_\mathbb{R}^*$ is the unique linear map that is given by the identity on $\ker(H_\alpha)$ (here $H_\alpha \in \mathfrak{t}_\mathbb{R} = (\mathfrak{t}_\mathbb{R}^*)^*$) and sends α to $-\alpha$. Suppose now that we choose an embedding $K \hookrightarrow \mathrm{U}(n)$. Then, as discussed in §7.12, we get an $\mathrm{Ad}(K)$-invariant inner product on $\mathfrak{g}$ given by $(x, y) \mapsto -\operatorname{trace}(x\theta(y))$ and hence a W-invariant inner product $\langle, \rangle$ on $\mathfrak{t}_\mathbb{R}$ given by $\langle x, y \rangle := \operatorname{trace}(xy)$. We may use this inner product to define a duality isomorphism $\mathfrak{t}_\mathbb{R}^* \cong \mathfrak{t}_\mathbb{R}$, $\lambda \mapsto u_\lambda$, where as before $\langle u_\lambda, v \rangle = \lambda(v)$ for $v \in \mathfrak{t}_\mathbb{R}$, and also an inner product $\langle, \rangle$ on $\mathfrak{t}_\mathbb{R}^*$ given by $\langle \lambda_1, \lambda_2 \rangle := \langle u_{\lambda_1}, u_{\lambda_2} \rangle$. Using such inner products, we may interpret the s_α geometrically as follows.

---

[16]End of half-lecture #15, Thursday, 11 Apr

Definition 123. Let $(V, \langle, \rangle)$ be a finite-dimensional real inner product space, and let $0 \neq v \in V$. The *reflection in v* is the linear map $r_v : V \to V$ that restricts to the identity on the orthogonal complement $v^\perp$ of v and that sends v to $-v$, thus

$$r_v(u) = u - 2\frac{\langle u, v \rangle}{\langle v, v \rangle} v.$$

Lemma 124. *Define inner products on $\mathfrak{t}_\mathbb{R}$ and $\mathfrak{t}_\mathbb{R}^*$ as above. Let $\alpha \in \Phi$. The action of s_α on $\mathfrak{t}_\mathbb{R}$ is the reflection in H_α, while that of s_α on $\mathfrak{t}_\mathbb{R}^*$ is the reflection in α.*

Proof. We recall from §7.13 that $H_\alpha = \frac{2}{\langle \alpha, \alpha \rangle} u_\alpha$, so that H_α and α correspond to positive multiples of one another under the isomorphism $\mathfrak{t}_\mathbb{R} \cong \mathfrak{t}_\mathbb{R}^*$ defined above. Thus $\ker(H_\alpha : \mathfrak{t}_\mathbb{R}^* \to \mathbb{R})$ is the orthogonal complement of α, while $\ker(\alpha : \mathfrak{t}_\mathbb{R} \to \mathbb{R})$ is the orthogonal complement of H_α. The required conclusion then follows from the formulas for s_α noted previously. $\square$

We've defined the regular subset $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ in $\mathfrak{t}_\mathbb{R}$ to be the complement of $\cup_{\alpha \in \Phi} \ker(\alpha)$, and we've defined a Weyl chamber C in $\mathfrak{t}_\mathbb{R}$ to be a connected component of $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$. We may similarly define the regular subset $\mathfrak{t}_\mathbb{R}^{*\,\mathrm{reg}}$ in $\mathfrak{t}_\mathbb{R}^*$ to be the complement of $\cup_{\alpha \in \Phi} \ker(H_\alpha)$, where now $H_\alpha \in \mathfrak{t}_\mathbb{R}$ is viewed as an element of $(\mathfrak{t}_\mathbb{R}^*)^*$, and a Weyl chamber $C^\vee \subseteq \mathfrak{t}_\mathbb{R}^*$ to be a connected component of $\mathfrak{t}_\mathbb{R}^{*\,\mathrm{reg}}$. Since the isometry $\mathfrak{t}_\mathbb{R} \cong \mathfrak{t}_\mathbb{R}^*$ defined above identifies H_α with a positive multiple of α, we see that it identifies the kernel of α with the kernel of H_α, and hence identifies each Weyl chamber C of $\mathfrak{t}_\mathbb{R}$ with a corresponding Weyl chamber $C^\vee$ of $\mathfrak{t}_\mathbb{R}^*$. This identification is independent of the embedding $K \hookrightarrow \mathrm{U}(n)$, since if C is described as

$$C = \{t \in \mathfrak{t}_\mathbb{R} : \varepsilon_\alpha \alpha(t) > 0 \text{ for all } \alpha \in \Phi\}$$

for some signs $\varepsilon_\alpha \in \{\pm 1\}$ (see §7.8), then $C^\vee$ is described analogously as

$$C^\vee = \{\lambda \in \mathfrak{t}_\mathbb{R}^* : \varepsilon_\alpha \lambda(H_\alpha) > 0 \text{ for all } \alpha \in \Phi\}.$$

In this way all of the facts that we prove below involving the Weyl group action on $\mathfrak{t}_\mathbb{R}$ and Weyl chambers thereof have immediate counterparts for $\mathfrak{t}_\mathbb{R}^*$.

Example 125. Take $K = \mathrm{SU}(3)$. Then (with the usual notation) $\Phi = \{\varepsilon_i - \varepsilon_j : 1 \leq i < j \leq 3\}$ and $\mathfrak{t}_\mathbb{R}^* \cong \mathfrak{t}_\mathbb{R} \cong (\mathbb{R}^3)_0$, the space of triples $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ with $\sum x_j = 0$. Using the standard representation, the inner product on $\mathfrak{t}_\mathbb{R}^*$ is given by the restriction of the standard inner product on $\mathbb{R}^3$. The elements

$$\varepsilon_1 - \varepsilon_2 = (1, -1, 0), \quad \varepsilon_2 - \varepsilon_3 = (0, 1, -1), \quad \varepsilon_3 - \varepsilon_1 = (-1, 0, 1)$$

of $\mathfrak{t}_\mathbb{R}^*$ sum to zero, have the same lengths, and have the same inner products with one another (up to sign). An isometric embedding of $\mathfrak{t}_{\mathbb{R}*}$ into $\mathbb{R}^2$ is thus obtained by sending these elements to the vertices of an equilateral triangle with center the origin. The following elements are then sent to the vertices of a regular hexagon with center the origin, traversed in the order they appear around the boundary:

$$\varepsilon_1 - \varepsilon_2, \quad \varepsilon_1 - \varepsilon_3, \quad \varepsilon_2 - \varepsilon_3, \quad \varepsilon_1 - \varepsilon_2, \quad \varepsilon_3 - \varepsilon_1, \quad \varepsilon_3 - \varepsilon_2.$$

In lecture we drew a picture and described the root reflections, noting that they preserve the set of roots. For each root α, we drew the hyperplane perpendicular to α, which we denoted by H_α. (The notation is not so terrible, because the hyperplane $H_\alpha \subseteq \mathfrak{t}_\mathbb{R}^*$ is also the kernel of the functional on $\mathfrak{t}_\mathbb{R}^*$ defined by the element $H_\alpha \in \mathfrak{t}_\mathbb{R}$.) There are six such hyperplanes. The connected components of their complements are the six Weyl chambers. Each Weyl chamber is a 60 degree open "pizza slice"

in $\mathbb{R}^2$. Each Weyl chamber contains exactly one root, which is positioned along the central ray of that Weyl chamber, 30 degrees from each of its walls. (TODO: replace this text with a picture.)

7.15. Using root reflections to elucidate the Weyl group.

Lemma 126. *Let C, C' be Weyl chambers in $\mathfrak{t}_{\mathbb{R}}$. Then there exist Weyl chambers $C = C_0, C_1, C_2, \ldots, C_{n-1}, C_n = C'$ and roots $\alpha_1, \ldots, \alpha_n \in \Phi$ with the following properties:*

  (i) *For $j = 1..n$, we have $\alpha_j > 0$ on $C_0, \ldots, C_{j-1}$ and $\alpha_j < 0$ on $C_j, \ldots, C_n$.*
  (ii) *Each $\beta \in \Phi - \{\pm\alpha_1, \ldots, \pm\alpha_n\}$ takes the same sign on $C_0, \ldots, C_n$.*
  (iii) *For $j = 1..n$, we have $s_{\alpha_j} C_{j-1} = C_j$. In particular,*

$$C' = s_{\alpha_n} \cdots s_{\alpha_1} C.$$

Analogous conclusions hold for any pair of Weyl chambers in $\mathfrak{t}_{\mathbb{R}}^*$.

The idea of the proof was illustrated in class on the picture of the Weyl chambers for SU(3). Consider the straight line path from some $t \in C$ to some $t' \in C'$. If t, t' are chosen generically, then this path will hit each root hyperplane $\alpha^{\perp} := \{x \in \mathfrak{t}_{\mathbb{R}} : \alpha(x) = 0\}$ one at a time. We take $\alpha_1, \ldots, \alpha_n$ to correspond to the root hyperplanes encountered along the way from t to t' and $C_0, \ldots, C_n$ the Weyl chambers, so that the path starts in the Weyl chamber C_0, then passes through the root hyperplane $\alpha_1^{\perp}$, then spends some time in the Weyl chamber C_1, then passes through the root hyperplane $\alpha_2^{\perp}$, and so on. We'll give the actual proof next time. For now, let's illustrate with some motivating applications.

Corollary 127. *The Weyl group $W = N(T)/T$ of a maximal torus T in a compact connected Lie group K is generated by the root reflections s_{α} ($\alpha \in \Phi = \Phi(K : T)$).*

Proof. Recall from Lemma 116 that W acts freely on the set of Weyl chambers. Let C be any Weyl chamber, and let $w \in W$. Then $C' := w(C)$ is a Weyl chamber. By Lemma 126, we may write $C' = s_{\alpha_n} \cdots s_{\alpha_1} C$ for some $\alpha_1, \ldots, \alpha_n \in \Phi$. Since the action of W on the set of Weyl chambers is free, we obtain $w = \alpha_n \cdots \alpha_1$. $\square$

Corollary 128. *With notation as assumptions as above, W acts simply-transitively on the set of Weyl chambers.*

Proof. We've seen (Lemma 116) that the action is defined and free; the final assertion in Lemma 126 implies that it is transitive. $\square$

[17]

Proof of Lemma 126. We implement the proof sketch given above. Fix $t' \in C'$. Set $\Sigma := \{\alpha \in \Phi : \alpha(C) > 0, \alpha(C') < 0\}$. Note that for each $t \in C$ and $\alpha \in \Phi$, we have $\alpha(t) > 0 > \alpha(t')$, so the line segment $\{(1 - \tau)t + \tau t' : \tau \in [0, 1]\}$ connecting t to t' intersects $\alpha^{\perp}$ in a unique point that we denote by $p_{\alpha}(t)$. This defines a map

$$p_{\alpha} : C \to \alpha^{\perp}.$$

This map is continuous and open. Moreover, for each $\alpha \in \Phi$, the hyperplane $\alpha^{\perp}$ is distinct from $\beta^{\perp}$ for all $\beta \in \Phi - \{\pm\alpha\} = \Phi - \mathbb{C}\alpha$, so the set

$$\alpha^{\perp} - \cup_{\beta \in \Phi - \{\pm\alpha\}} \alpha^{\perp} \cup \beta^{\perp}$$

---

[17]End of half-lecture #16, Tuesday, 16 Apr

is dense and open in $\alpha^\perp$. By an exercise in topology, it follows that for each $\alpha \in \Sigma$, the set
$$A_\alpha := \{t \in C : p_\alpha(t) \in \beta^\perp \text{ for all } \beta \in \Phi - \mathbb{C}\alpha\}$$
is dense and open in C. Hence likewise $A := \cap_{\alpha \in \Sigma} A_\alpha$ is dense and open in C. Choose any $t \in A$. Let $\gamma : [0,1] \to \mathfrak{t}_\mathbb{R}$ denote the linear path connecting t to t', thus $\gamma(\tau) := (1-\tau)t + \tau t'$. For each $\alpha \in \Sigma$, abbreviate $p_\alpha := p_\alpha(t)$, and let $\tau_\alpha \in (0,1)$ be such that $\gamma(\tau_\alpha) = p_\alpha$. By the construction of t, we have $\tau_\alpha \neq \tau_{\alpha'}$ whenever $\alpha \neq \alpha'$. We may thus order $\Sigma = \{\alpha_1, \ldots, \alpha_n\}$ in such a way that $\tau_{\alpha_1} < \cdots < \tau_{\alpha_n}$. Abbreviate $\tau_j := \tau_{\alpha_j}$ and set $\tau_0 := 0, \tau_{n+1} := 1$. Note if $\beta \in \Phi - (\Sigma \cup (-\Sigma))$, then β takes the same sign on t and t', hence also on the path γ connecting them. It follows that for $j = 0..n$, the set $\{\gamma(\tau) : \tau_j < \tau < \tau_{j+1}\}$ belongs to some Weyl chamber C_j of $\mathfrak{t}_\mathbb{R}$, with $C_0 = C$ and $C_n = C'$. Moreover, for $j = 1..n$, there is a neighborhood U_j of p_{α_j} that doesn't intersect $\beta^\perp$ for any $\beta \in \Phi - \mathbb{C}\alpha_j$, thus $U_j = (U_j \cap C_{j-1}) \sqcup (U_j \cap \alpha_j^\perp) \sqcup (U_j \cap C_j)$. The root reflection s_{α_j} satisfies $s_{\alpha_j}|_{\alpha_j^\perp} = \mathrm{id}$, thus $s_{\alpha_j} U_j \cap U_j \neq \emptyset$, thus $s_{\alpha_j}(U_j \cap C_{j-1}) \cap (U_j \cap C_{j-1}) \neq \emptyset$, hence $s_{\alpha_j} C_{j-1} = C_j$. The remaining assertions follow from what we have already shown. $\square$

Corollary 129. *Let C be a Weyl chamber. Any $x \in \mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ is W-conjugate to a unique element of C. The W-stabilizer of x is trivial.*

Proof. We use that x belongs to some Weyl chamber and that W acts transitively on the set of Weyl chambers. $\square$

Corollary 130. *Let C be a Weyl chamber. Then any $x \in \mathfrak{t}_\mathbb{R}$ is W-conjugate to a unique element of the closure $\overline{C}$ of C.*
    The W-stabilizer of x is generated by $\{s_\alpha : \alpha(x) = 0\}$.

Proof. We observe first that $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ is dense in $\mathfrak{t}_\mathbb{R}$. Indeed, if $x \in \mathfrak{t}_\mathbb{R}$ and y is any element of $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$, then $x + \varepsilon y \in \mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ for all small enough $\varepsilon > 0$, and $x = \lim_{\varepsilon \to 0}(x + \varepsilon y)$. Since $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ is the union of the W-conjugates of C, it follows that $\mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$ is the union of the W-conjugates of $\overline{C}$. This gives the existence of a W-conjugate of x in $\overline{C}$. For verifying the uniqueness, we may assume that x belongs to $\overline{C}$, and must show that if $w \in W$ satisfies $w(x) \in \overline{C}$, then $w(x) = x$. To see this, we apply Lemma 126 to C and $C' := w^{-1}(C)$, giving us some chambers $C_0, \ldots, C_n$ and roots $\alpha_1, \ldots, \alpha_n$. Our hypothesis implies that $x \in \overline{C} \cap \overline{C'}$. For $j = 1..n$, we have $\alpha_j > 0$ on C and $\alpha_j < 0$ on C', thus $\alpha_j(x) = 0$, and so $s_{\alpha_j}(x) = x$. On the other hand, $w^{-1} = s_{\alpha_n} \cdots s_{\alpha_1}$, so $w = s_{\alpha_1} \cdots s_{\alpha_n}$. Thus $w(x) = x$.
    The final assertion follows from the argument just given; we note that it doesn't depend upon the choice of C. $\square$

7.16. Basics on dominance, positivity and simple roots.

Definition 131. Choose a Weyl chamber $C \subseteq \mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$, hence a corresponding dual Weyl chamber $C^\vee \subseteq \mathfrak{t}_\mathbb{R}^{\mathrm{reg}}$. We say that $x \in \mathfrak{t}_\mathbb{R}$ (resp. $\lambda \in \mathfrak{t}_\mathbb{R}^*$) is *dominant* if $x \in \overline{C}$ (resp. if $\lambda \in \overline{C^\vee}$), and *strictly dominant* if $x \in C$ (resp. if $\lambda \in C^\vee$).
    We say that $\alpha \in \Phi$ is *positive* if $\alpha(C) > 0$; we abbreviate this condition simply to $\alpha > 0$ when C is clear by context. We denote by $\Phi^+ := \Phi^+(C) \subseteq \Phi$ the subset of positive roots. We say that a positive root $\alpha \in \Phi^+$ is *simple* if it cannot be written in the form $\alpha_1 + \alpha_2$ for some $\alpha_1, \alpha_2 \in \Phi^+$. We denote by $\Delta := \Delta(C) \subseteq \Phi^+$ the set of simple roots.

For $x, y \in \mathfrak{t}_\mathbb{R}$ we say that x is *higher than* y, and write $x \geq y$, if $\lambda(x - y) \geq 0$ for all $\lambda \in C^\vee$. Similarly, for $\lambda, \mu \mathfrak{t}_\mathbb{R}^*$, we say that λ is *higher than* μ, and write $\lambda \geq \mu$, if $(\lambda - \mu)(x) \geq 0$ for all $x \in C$. We use "strictly higher than" to mean "higher than and not equal to."

We note that, for reasons explained earlier, all of these notions are compatible with the isometry $\mathfrak{t}_\mathbb{R} \cong \mathfrak{t}_\mathbb{R}^*$ defined by the inner product given by the trace pairing with respect to any embedding $K \hookrightarrow \mathrm{U}(n)$.

Lemma 132. *(Not proved in lecture, but useful facts to mention at this point.)*
  (i) $\Phi = \Phi^+ \sqcup (-\Phi^+)$.
  (ii) $C = \{x \in \mathfrak{t}_\mathbb{R} : \alpha(x) > 0 \text{ for all } \alpha > 0\}$ and $C^\vee = \{\lambda \in \mathfrak{t}_\mathbb{R}^* : \lambda(H_\alpha) > 0 \text{ for all } \alpha > 0\}$.
  (iii) $\{x \in \mathfrak{t}_\mathbb{R} : x \geq 0\} = \sum_{\alpha > 0} \mathbb{R}_{\geq 0} H_\alpha$ and $\{\lambda \in \mathfrak{t}_\mathbb{R}^* : \lambda \geq 0\} = \sum_{\alpha > 0} \mathbb{R}_{\geq 0} \alpha$.

Proof. We've seen that we may write $C = \{x : \varepsilon_\alpha \alpha(x) > 0 \text{ for all } \alpha \in \Phi\}$ for some signs $\varepsilon_\alpha \in \{\pm 1\}$. Then $C^\vee = \{\lambda : \varepsilon_\alpha \lambda(H_\alpha) > 0 \text{ for all } \alpha \in \Phi\}$. Since $\varepsilon_{-\alpha} = -\varepsilon_\alpha$ and $\Phi^+ = \{\alpha : \varepsilon_\alpha = 1\}$, we deduce assertions (i) and (ii). The final assertion (iii) boils down, as follows, to a general duality principle for convex cones. Let M_1 and M_2 denote the LHS and RHS of the claimed identity. Then M_1 and M_2 are closed convex cones containing the origin. Clearly $M_2 \subseteq M_1$. To prove the reverse containment $M_1 \subseteq M_2$, it suffices (by the separating hyperplane theorem, the finite-dimensional case of Hahn–Banach) to show that for each $x \in \mathfrak{t}_\mathbb{R}$ with $M_2(x) \geq 0$, we have $M_1(x) \geq 0$. Indeed, if $M_2(x) \geq 0$, then $\alpha(x) \geq 0$ for all $\alpha > 0$; since $\lambda \geq 0$ for all $\lambda \in M_1$, the required implication follows. $\square$

For instance, take $K = \mathrm{U}(n)$, with T as usual and coordinates as before. We might take

$$C = \{x \in \mathfrak{t}_\mathbb{R} \cong \mathbb{R}^n : x_1 > \cdots > x_n\}. \tag{71}$$

Then

$$C^\vee = \{\lambda \in \mathfrak{t}_\mathbb{R}^* \cong \mathbb{R}^n : \lambda_1 > \cdots > \lambda_n\}$$

and

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j : i < j\} = \{\varepsilon_1 - \varepsilon_2, \varepsilon_1 - \varepsilon_3, \dots\} \tag{72}$$

and

$$\Delta = \{\varepsilon_1 - \varepsilon_2, \varepsilon_2 - \varepsilon_3, \dots, \varepsilon_{n-1} - \varepsilon_n\}.$$

Corollary 133. *If $\lambda \in \mathfrak{t}_\mathbb{R}^*$ is dominant, then $\lambda \geq w(\lambda)$ for all $w \in W$.*

Proof. We apply Lemma 126 in its dual form to $C^\vee$ and $wC^\vee$. This gives us a sequence of dual Weyl chambers $C_0^\vee = C^\vee, C_1^\vee, \dots, C_n^\vee = wC^\vee$ and roots $\alpha_1, \dots, \alpha_n$ satisfying (among other properties) $H_{\alpha_j} > 0$ on $C_0^\vee, \dots, C_{j-1}^\vee$ and $H_{\alpha_j} < 0$ on $C_j^\vee, \dots, C_n^\vee$. Set $\lambda_j := s_{\alpha_j} \cdots s_{\alpha_1} \lambda$. Then $\lambda = \lambda_0$ and $w(\lambda) = \lambda_n$, so

$$\lambda - w(\lambda) = (\lambda_0 - \lambda_1) + (\lambda_1 - \lambda_2) + \cdots + (\lambda_{n-1} - \lambda_n).$$

For $j = 1..n$, the formulas defining λ_j and s_{α_j} give

$$\lambda_{j-1} - \lambda_j = \lambda_{j-1} - s_{\alpha_j} \lambda_{j-1} = \lambda_{j-1}(H_{\alpha_j}) \alpha_j.$$

We have $\alpha_j > 0$ (because $H_{\alpha_j} > 0$ on $C^\vee$) and $\lambda_{j-1}(H_{\alpha_j}) \geq 0$ (because $H_{\alpha_j} \geq 0$ on C_{j-1}), hence $\lambda_{j-1} - \lambda_j \geq 0$. It follows that $\lambda - w(\lambda) \geq 0$. $\square$

Theorem 134. Δ *is linearly independent over* $\mathbb{R}$, *and the* $\mathbb{Z}$-*module* $\mathbb{Z}\Phi$ *spanned by* Φ *admits the* $\mathbb{Z}$-*module basis* Δ.

We'll give the proof after a few lemmas.

Lemma 135. *Every* $\beta \in \Phi^+$ *may be written in the form* $\beta = \sum_{\alpha \in \Delta} n_\alpha \alpha$ *with* $n_\alpha \in \mathbb{Z}_{\geq 0}$.

Proof. Fix $t \in C$. As β varies over Φ^+, the quantities $\beta(t)$ vary over a finite subset of $\mathbb{R}_{>0}$. We argue by induction on $\beta(t)$. If $\beta \in \Delta$, then we are done. Otherwise $\beta = \beta_1 + \beta_2$ with $\beta_1, \beta_2 \in \Delta$. Then $\beta_1(t), \beta_2(t) < \beta(t)$, so we may conclude by our inductive hypothesis. $\square$

Let $\alpha, \beta \in \Phi$ be roots. Recall that, if we choose an embedding $K \hookrightarrow \mathrm{U}(n)$ and hence an inner product $\langle x, y \rangle = \mathrm{trace}(xy)$ on $\mathfrak{t}_\mathbb{R}$ and hence an isometry $\mathfrak{t}_\mathbb{R}^* \ni \lambda \mapsto u_\lambda \in \mathfrak{t}_\mathbb{R}$, then $H_\alpha = \frac{2}{(\alpha,\alpha)} u_\alpha$. Thus

$$\langle \alpha, \beta \rangle = \beta(u_\alpha) = \alpha(u_\beta) = \frac{\langle \alpha, \alpha \rangle}{2} \beta(H_\alpha) = \frac{\langle \beta, \beta \rangle}{2} \alpha(H_\beta). \tag{73}$$

In particular, the quantities

$$(\alpha, \beta), \quad \alpha(H_\beta), \quad \beta(H_\alpha)$$

all have the same sign. We say that the roots α, β

- are *orthogonal* if these quantities are zero,
- form an *acute* angle if these quantities are positive, and
- form an *obtuse* angle if they are negative.

These notions are evidently independent of the choice of embedding $K \hookrightarrow \mathrm{U}(n)$.

Lemma 136. *Let* $\alpha, \beta \in \Phi$ *be roots that are non-proportional (thus* $\alpha \neq \pm\beta$*).*

- *If* α, β *are orthogonal, then* $\alpha(H_\beta) = \beta(H_\alpha) = 0$.
- *Suppose* α, β *form an acute angle. Then the pair* $(\alpha(H_\beta), \beta(H_\alpha))$ *is one of the following:*
$$(1,1), (1,2), (2,1), (1,3), (3,1).$$
  Moreover, $\alpha - \beta$ *is a root.*
- *Suppose* α, β *form an obtuse angle. Then the pair* $(\alpha(H_\beta), \beta(H_\alpha))$ *is one of the following:*
$$(-1,-1), (-1,-2), (-2,-1), (-1,-3), (-3,-1).$$
  Moreover, $\alpha + \beta$ *is a root.*

Proof. Suppose for instance that α, β form an acute angle. Since $\alpha, \beta \in \mathfrak{t}_\mathbb{Z}^*$ and $H_\alpha, H_\beta \in \mathfrak{t}_\mathbb{Z}$, we then have $\alpha(H_\beta), \beta(H_\alpha) \in \mathbb{Z}_{\geq 1}$. Since α, β are non-proportional, we may apply the Cauchy–Schwartz inequality in its strict form and the identity (73) to see that

$$\alpha(H_\beta)\beta(H_\alpha) = 4\frac{|\langle \alpha, \beta \rangle|^2}{\langle \alpha, \alpha \rangle \langle \beta, \beta \rangle} < 4.$$

Thus $\alpha(H_\beta)\beta(H_\alpha) \in \{1, 2, 3\}$, which leads to the possibilities indicated. In particular, either $\alpha(H_\beta)$ or $\beta(H_\alpha)$ is equal to 1. Suppose for instance that $\beta(H_\alpha) = 1$. Then

$$s_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha = \beta - \alpha,$$

so $\beta - \alpha$ is a root, thus $\alpha - \beta = -(\beta - \alpha)$ is a root, as required. The case that $\alpha(H_\beta) = 1$ is treated similarly. The obtuse case is treated similarly. $\square$

Corollary 137. *If* α, β *are distinct elements of* Δ, *then* $\langle \alpha, \beta \rangle \leq 0$.

Proof. We've seen that otherwise $\gamma := \alpha - \beta$ is a root. If $\gamma \in \Phi^+$, then the identity $\alpha = \beta + \gamma$ implies that $\alpha \notin \Delta$, while if $-\gamma \in \Phi^+$, then the identity $\beta = \alpha + (-\gamma)$ implies that $\beta \notin \Delta$. In either case we obtain the required contradiction. $\qquad \square$

Proof of Theorem 134. Suppose there are distinct elements $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in \Delta$ and positive reals $a_1, \ldots, a_m, b_1, \ldots, b_n$ so that

$$x := \sum a_i \alpha_i = \sum b_j \beta_j =: y.$$

Corollary 137 then implies that $\langle \alpha_i, \beta_j \rangle \leq 0$, so

$$0 \leq \langle x, x \rangle = \langle x, y \rangle = \sum_{i,j} \underbrace{a_i b_j}_{>0} \underbrace{\langle \alpha_i, \beta_j \rangle}_{\leq 0} \leq 0,$$

so equality holds, and thus $x = y = 0$. Choose any $t \in C$. Then $a_i(t) > 0$, so $0 = x(t) = \sum_i a_i \alpha_i(t)$ and $0 = y(t) = \sum_j b_j \beta_j(t)$ with each summand positive. Thus $m = n = 0$. $\qquad \square$

Theorem 138. *The Weyl group is generated by the simple reflections, i.e., $W = \langle s_\alpha : \alpha \in \Delta \rangle$.*

Proof. We'll leave this to the homework. It can be proved using Lemma 126; the main point is that $\alpha_1 \in \Delta(C)$ for any α_1 in the conclusion of that lemma. $\qquad \square$

7.17. **Cartan matrices.** Let K be a compact connected Lie group, with maximal torus T, Weyl chamber C and accompanying notation as above.

Definition 139. The Cartan matrix is $(N_{\alpha\beta})_{\alpha,\beta \in \Delta}$ where $N_{\alpha\beta} := \alpha(H_\beta)$.

We'll note that the Cartan matrix, as we've defined it here, depends only upon the isomorphism class of the compact Lie group K. (Use that any two maximal tori are G-conjugate and that any two Weyl chambers are W-conjugate.)

Lemma 140. *We have $N_{\alpha\alpha} = 2$. For $\alpha \neq \beta$, the pair $(N_{\alpha\beta}, N_{\beta\alpha})$ is one of*

$$(0,0), \quad (-1,-1), \quad (-1,-2), \quad (-2,-1), \quad (-1,-3), \quad (-3,-1).$$

Proof. By Lemma 136 and Corollary 137. $\qquad \square$

We'll prove the following (and some variants) after the break:

Theorem 141. *The tuple $(T, \Delta, \Delta^\vee)$ determines K up to isomorphism, in the following sense:*

Suppose given two compact connected Lie groups K', K'' equipped with maximal tori T', T'', Weyl chambers C', C''. Let $\Delta' = \{\gamma_1', \ldots, \gamma_n'\}$ and $\Delta'' = \{\gamma_1'', \ldots, \gamma_n''\}$ denote the associated sets of simple roots, and $H_i' := H_{\gamma_i'}$, $H_i'' := H_{\gamma_i''}$ the simple coroots. Suppose given an isomorphism $\kappa : T' \to T''$ such that $\kappa^* \gamma_i'' = \gamma_i'$ and $\kappa_* H_i' = H_i''$. Then κ extends to an isomorphism $K' \to K''$.

More precisely, suppose given maps $F_i' : \mathrm{SU}(2) \to K'$ and $F_i'' : \mathrm{SU}(2) \to K''$ attached to the roots γ_i' and γ_i''. Let X_i', Y_i', H_i' and X_i'', Y_i'', H_i'' be defined accordingly. Then there exists a unique extension of κ whose differential maps X_i' to X_i'' for each i.

This has the following consequence:

Corollary 142. *Suppose that K has finite center, or equivalently, that $\mathfrak{k}$ has trivial center. Then the pair $(T, (N_{\alpha\beta})_{\alpha,\beta \in \Delta})$ determines K up to isomorphism in the following sense:*

Let K', K'' be two compact connected Lie groups having finite center, with maximal tori T', T'' and Weyl chambers C', C''' and sets of simple roots Δ', Δ'' so that there is an isomorphism $T' \xrightarrow{\simeq} T''$ for which the induced isomorphism $X(T'') \xrightarrow{\simeq} X(T')$ defines a bijection $\Delta'' \xrightarrow{\simeq} \Delta'$ with respect to which the Cartan matrices coincide, i.e., $\alpha''(H_{\beta''}) = \alpha'(H_{\beta'})$ whenever $\alpha', \beta' \in \Delta'$ correspond to $\alpha'', \beta'' \in \Delta''$. Then the isomorphism $T' \xrightarrow{\simeq} T''$ extends to an isomorphism $K' \xrightarrow{\simeq} K''$. The set of possible isomorphisms may be described as in theorem 141.

[18] The proof will be given in §??? after some preliminaries.

7.18. Strings of roots.

Let K, T, Φ be as usual: a compact connected Lie group, a maximal torus, and the set of roots. Let $\alpha, \beta \in \Phi$.

Lemma 143. (i) *If $\alpha + \beta \in \Phi$, then $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.*
(ii) *Suppose that α, β are non-proportional, i.e., $\mathbb{Q}\alpha \neq \mathbb{Q}\beta$. The set*

$$E := \{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\}$$

contains 0 and is finite, hence has a minimal element of the form $-p \in \mathbb{Z}_{\leq 0}$ and a maximal element of the form $q \in \mathbb{Z}_{\geq 0}$. We have

$$E = \mathbb{Z} \cap [-p, q]. \tag{74}$$

Moreover, for any nonzero elements $X_\alpha \in \mathfrak{g}^\alpha, Y_\alpha \in \mathfrak{g}^{-\alpha}$, each of the maps

$$\mathfrak{g}^{\beta-p\alpha} \xrightarrow{\mathrm{ad}_{X_\alpha}} \mathfrak{g}^{\beta-(p-1)\alpha} \xrightarrow{\mathrm{ad}_{X_\alpha}} \cdots \xrightarrow{\mathrm{ad}_{X_\alpha}} \mathfrak{g}^{\beta+(q-1)\alpha} \xrightarrow{\mathrm{ad}_{X_\alpha}} \mathfrak{g}^{\beta+q\alpha},$$

$$\mathfrak{g}^{\beta-p\alpha} \xleftarrow{\mathrm{ad}_{Y_\alpha}} \mathfrak{g}^{\beta-(p-1)\alpha} \xleftarrow{\mathrm{ad}_{Y_\alpha}} \cdots \xleftarrow{\mathrm{ad}_{Y_\alpha}} \mathfrak{g}^{\beta+(q-1)\alpha} \xleftarrow{\mathrm{ad}_{Y_\alpha}} \mathfrak{g}^{\beta+q\alpha},$$

is an isomorphism. Moreover,

$$\beta(H_\alpha) = p - q. \tag{75}$$

Proof. Note first that (ii) implies (i): if $\alpha+\beta \in \Phi$, then (since $\mathbb{Q}\alpha \cap \Phi = \{\pm\alpha\}$) α, β are non-proportional, the set E as defined in (ii) contains the elements $0, 1$, and $\mathrm{ad}(X_\alpha) : \mathfrak{g}^\beta \to \mathfrak{g}^{\beta+\alpha}$ is an isomorphism of one-dimensional vector spaces, which implies that $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.

We turn to the proof of (ii). Since β, α are non-proportional, we have $\beta+k\alpha \neq 0$ for all $k \in \mathbb{Z}$. Thus

$$V := \oplus_{k \in \mathbb{Z}} \mathfrak{g}^{\beta+k\alpha} = \oplus_{k \in E} \mathfrak{g}^{\beta+k\alpha}. \tag{76}$$

We check easily that V is an $\mathfrak{sl}_2(\mathbb{C})_\alpha$-module whose weight spaces, i.e., $\mathrm{ad}(H_\alpha)$-eigenspaces, are the summands $\mathfrak{g}^{\beta+k\alpha}$, with correspond eigenvalues $(\beta+k\alpha)(H_\alpha) = \beta(H_\alpha) + 2k$. The weight spaces are thus one-dimensional and the weights all have the same parity, so by the structure theory for $\mathfrak{sl}_2(\mathbb{C})$-modules, we deduce that V is irreducible, that the weights of V are of the form $\{-n, -n+2, \ldots, n-2, n\}$ for some nonnegative integer n, that (74) holds, and that the connecting maps between adjacent weight spaces defined by $\mathrm{ad}(X_\alpha)$ and $\mathrm{ad}(Y_\alpha)$ are isomorphisms. Moreover, $n = \beta(H_\alpha) + 2q$ and $-n = \beta(H_\alpha) - 2p$, which gives (75). $\qquad\square$

---

[18]End of lecture #17, Thursday, 18 Apr

7.19. **Generating the Lie algebra using simple root vectors.** For the next few results, let $K, T, C, \Phi, \Phi^+, \Delta$ be as usual. Write

$$\Delta = \{\gamma_1, \ldots, \gamma_n\}.$$

Fix $F_{\gamma_j} : \mathrm{SU}(2) \to K$ as in §7.10 and associated elements $X_j := X_{\gamma_j}$, $Y_j := Y_{\gamma_j}$, $H_j := H_{\gamma_j}$.

Lemma 144. Let $\beta \in \Phi^+$. Then there exist $\alpha_1, \ldots, \alpha_r \in \Delta$ so that

- $\beta = \alpha_1 + \cdots + \alpha_r$, *and*
- $\alpha_1 + \cdots + \alpha_s \in \Phi$ *for $s = 1..r$.*

Proof. The proof is an adaptation of that of Lemma 135. Fix $t \in C$. Induct on $\beta(t) \in \mathbb{R}_{>0}$. If $\beta \in \Delta$, then the required conclusion holds with $r = 1$ and $\alpha_1 = \beta$, so suppose $\beta \notin \Delta$.

We claim that there exists $j \in \{1..n\}$ so that β and γ_j form an acute angle. Otherwise, arguing as in the proof of Theorem 134, we see that the set $\{\beta, \gamma_1, \ldots, \gamma_n\}$ is linearly independent over $\mathbb{R}$. Since $\beta \in \oplus_{j=1..n}\mathbb{Z}_{\geq 0}\gamma_j$, we obtain the required contradiction.

Fix one such j. Since $\beta \in \Phi^+ - \Delta$, the roots β and γ_j are non-proportional, so by Lemma 136, the difference $\beta - \gamma_j$ is a root. We claim that $\beta - \gamma_j$ is positive. If not, then $\gamma_j - \beta$ is positive. Writing $\beta = \sum_i c_i \gamma_i$ with $c_i \in \mathbb{Z}_{\geq 0}$, the decomposition $\gamma_j - \beta = \sum_i (\delta_{ij} - c_i)\gamma_i$ then has nonnegative coefficients, so $c_j \leq 1$ and $c_i \leq 0$ for $i \neq j$. These conditions force either $\beta = 0$ (which is impossible because β is a root) or $\beta = \gamma_j$ (which is impossible because $\beta \notin \Delta$).

We now apply our inductive hypothesis to $\beta - \gamma_j$, writing it as $\alpha_1 + \cdots + \alpha_r$ for some simple roots satisfying the required conclusion of the lemma, and take $\alpha_{r+1} := \gamma_j$. Then $\beta = \alpha_1 + \cdots + \alpha_{r+1}$ gives the required decomposition. $\square$

Corollary 145. $\mathfrak{g}$ *is generated (as a complex Lie algebra) by $\mathfrak{t}_{\mathbb{C}} \cup \{X_i\} \cup \{Y_j\}$. More precisely,*

$$\mathfrak{g} = \mathfrak{t}_{\mathbb{C}} \oplus \Big(\sum_{\substack{r \geq 1 \\ i=(i_1,\ldots,i_r)\in\mathbb{Z}_{\geq 1}^r}} \mathbb{C}e_i \Big) \oplus \Big(\sum_{\substack{r \geq 1 \\ i=(i_1,\ldots,i_r)\in\mathbb{Z}_{\geq 1}^r}} \mathbb{C}f_i \Big), \tag{77}$$

where

$$e_i := [X_{i_1}, \ldots, [X_{i_{r-1}}, X_{i_r}] \cdots] \in \mathfrak{g}^{\gamma_{i_1} + \cdots + \gamma_{i_r}}, \tag{78}$$

$$f_i := [Y_{i_1}, \ldots, [Y_{i_{r-1}}, Y_{i_r}] \cdots] \in \mathfrak{g}^{-\gamma_{i_1} - \cdots - \gamma_{i_r}}. \tag{79}$$

The action of the indicated generators for $\mathfrak{g}$ on the above decomposition is described as follows: For $H \in \mathfrak{t}_{\mathbb{C}}$, we have

$$\mathrm{ad}_{X_j} H = -\gamma_j(H)X_j,$$

$$\mathrm{ad}_{Y_j} H = -\gamma_j(H)X_j,$$

$$\mathrm{ad}_{X_j} e_i = [X_j, [X_{i_1}, \ldots, [X_{i_{r-1}}, X_{i_r}] \cdots]]$$

$$\mathrm{ad}_H e_i = \langle H, \gamma_{i_1} + \cdots + \gamma_{i_r} \rangle e_i$$

$$\mathrm{ad}_{Y_j} f_i = [Y_j, [Y_{i_1}, \ldots, [Y_{i_{r-1}}, Y_{i_r}] \cdots]]$$

$$\mathrm{ad}_H f_i = -\langle H, \gamma_{i_1} + \cdots + \gamma_{i_r} \rangle f_i$$

$$\mathrm{ad}_{Y_j} X_i = \begin{cases} -H_i & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

$$\mathrm{ad}_{Y_j} e_i = [\mathrm{ad}_{Y_j} X_{i_1}, \ldots, [X_{i_{r-1}}, X_{i_r}] \cdots] + \cdots + [X_{i_1}, \ldots, [X_{i_{r-1}}, \mathrm{ad}_{Y_j} X_{i_r}] \cdots],$$

$$\mathrm{ad}_{X_j} Y_i = \begin{cases} H_i & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

$$\mathrm{ad}_{X_j} f_i = [\mathrm{ad}_{X_j} Y_{i_1}, \ldots, [Y_{i_{r-1}}, Y_{i_r}] \cdots] + \cdots + [Y_{i_1}, \ldots, [Y_{i_{r-1}}, \mathrm{ad}_{X_j} Y_{i_r}] \cdots].$$

Proof. Let $\mathfrak{h}$ denote the subalgebra of $\mathfrak{g}$ generated by the indicated set. In particular, $\mathfrak{h}$ is a $\mathfrak{t}_{\mathbb{C}}$-module containing $\mathfrak{g}^{\pm \gamma_j}$ for $j = 1..n$, so we may write

$$\mathfrak{h} = \mathfrak{t}_{\mathbb{C}} \oplus (\oplus_{\alpha \in E} \mathfrak{g}^{\alpha})$$

for some subset $E \subseteq \Phi$ containing $\{\pm \gamma_1, \ldots, \pm \gamma_n\}$. We must show that in fact $E = \Phi$. So let $\beta \in \Phi$; we must show that $\beta \in E$. We assume that $\beta \in \Phi^+$; a similar proof applies if $-\beta \in \Phi^+$. By Lemma 144, we may write $\beta = \alpha_1 + \cdots + \alpha_r$ with each partial sum $\alpha_1 + \cdots + \alpha_s$ also a root. We show by induction on r that $\beta \in E$. If $r = 1$, then $\beta \in \Delta \subseteq E$, so suppose $r \geq 2$. By our inductive hypothesis, $\delta := \alpha_1 + \cdots + \alpha_{r-1} \in E$. We have $\beta = \delta + \alpha_r$, with each term a root, so by Lemma 143, $\mathfrak{g}^{\beta} = [\mathfrak{g}^{\delta}, \mathfrak{g}^{\alpha_r}]$. Our inductive hypothesis gives $\mathfrak{g}^{\delta}, \mathfrak{g}^{\alpha_r} \subseteq \mathfrak{h}$. Since $\mathfrak{h}$ is a subalgebra, it follows that $\mathfrak{g}^{\beta} \subseteq \mathfrak{h}$, whence as required that $\beta \in E$.

This establishes that the indicated set is indeed a generating set for $\mathfrak{g}$. The remaining assertions follow from the argument just given together with some routine calculation. $\square$

Corollary 146. *Any nonzero ideal $\mathfrak{a}$ of $\mathfrak{g}$ contains a nonzero element of $\mathfrak{t}_{\mathbb{C}}$.*

Proof. The proof is similar to that of the previous corollary (and could even be deduced from its conclusion by considering orthogonal complements). Since $\mathfrak{a}$ is in particular a $\mathfrak{t}_{\mathbb{C}}$-module, we may write $\mathfrak{a} = \mathfrak{b} \oplus (\oplus_{\alpha \in E} \mathfrak{g}^{\alpha})$ for some subspace $\mathfrak{b} \subseteq \mathfrak{t}_{\mathbb{C}}$ and some subset $E \subseteq \Phi$. We want to show that $\mathfrak{b}$ is nonzero. Suppose otherwise that $\mathfrak{b}$ is zero. Then E is nonempty. Let $\beta \in E$. Write $\beta = \alpha_1 + \cdots + \alpha_r$ as in Lemma 144. Set $Z := [Y_{\alpha_1}, \cdots, [Y_{\alpha_{r-1}}, [Y_{\alpha_r}, X_{\beta}]] \cdots]$. Then $Z \in \mathfrak{t}_{\mathbb{C}} \cap \mathfrak{b} = \mathfrak{b} = \{0\}$, but several applications of Lemma 143 (and (66)) implies that $Z \neq 0$, giving the required contradiction. $\square$

7.20. **Proof of Theorem 141.** We need one final miscellaneous lemma:

Lemma 147. *Let K be a compact connected Lie group with maximal torus T. Then the map $\pi_1(T) \to \pi_1(K)$ induced by the inclusion $T \to K$ is surjective.*

Proof. (Omitted in lecture; see for instance Section 36.4 of my Fall 2016 notes on Lie groups, linked on the course homepage.) $\square$

We now give the proof of Theorem 141. Let notation be as in its statement. We may write $\Delta' = \{\gamma_1', \ldots, \gamma_n'\}$ and $\Delta'' = \{\gamma_1'', \ldots, \gamma_n''\}$ so that γ_i' corresponds to γ_i''. We define $\mathfrak{g}', \mathfrak{g}''$ in the obvious way and define $X_i', Y_j' \in \mathfrak{g}'$ and $X_i'', Y_j'' \in \mathfrak{g}''$ as in §7.19. We will show more precisely that there exists a unique isomorphism $K' \xrightarrow{\cong} K''$ extending the given isomorphism $\kappa : T' \xrightarrow{\cong} T''$ with the additional property that the induced isomorphism $\mathfrak{g}' \xrightarrow{\cong} \mathfrak{g}''$ maps X_i' to X_i'' for $i = 1..n$.

To that end, we fix linear embeddings of K' and K'', hence involutions θ of $\mathfrak{g}'$ and $\mathfrak{g}''$ with fixed subspaces $\mathfrak{k}'$ and $\mathfrak{k}''$. It suffices to show that there exists a unique θ-equivariant isomorphism $\mathfrak{g}' \xrightarrow{\cong} \mathfrak{g}''$ given on $\mathfrak{t}_{\mathbb{C}}'$ by $d\kappa$ and which maps X_i' to X_i''. Indeed, such an isomorphism of complexified Lie algebras induces an isomorphism $\tau : \mathfrak{k}' \xrightarrow{\cong} \mathfrak{k}''$ of their θ-fixed subspaces. What remains to be checked is just that any such isomorphism arises from a Lie group morphism $K' \to K''$, and

similarly for the inverse isomorphism. From basic Lie theory (see e.g. Theorem 155 of the Fall 2016 notes linked on the course webpage), the Lie group K' admits a universal connected covering group $\widetilde{K'}$, which fits into a short exact sequence $1 \to \pi_1(K) \to \widetilde{K'} \to K' \to 1$, where $\pi_1(K)$ identifies with a discrete central subgroup of $\widetilde{K'}$. The given Lie algebra map $\mathfrak{k}' \to \mathfrak{k}''$ extends to a Lie group map $\widetilde{K'} \to K''$ (see e.g. Theorem 146 of the aforementioned notes); we just need to check that $\pi_1(K')$ lies in the kernel of the latter. But by Lemma 147, it suffices to show that $\pi_1(T')$ lies in the indicated kernel, which follows from the fact that the given map $T' \to T''$ is a well-defined Lie group map. The completes the proof of the required reduction.

It remains to produce the required θ-equivariant map $\mathfrak{g}' \to \mathfrak{g}''$. We will do so by constructing its graph. Let $\mathfrak{g}$ denote the subalgebra of $\mathfrak{g}' \oplus \mathfrak{g}''$ generated by $\mathfrak{t}_{\mathbb{C}} \cup \{X_i\} \cup \{Y_i\}$, where

$$\mathfrak{t}_{\mathbb{C}} := \{(x', x'') \in \mathfrak{t}'_{\mathbb{C}} \oplus \mathfrak{t}''_{\mathbb{C}} : d\kappa(x') = x''\}$$

and $X_i := (X'_i, X''_i)$ and $Y_i := (Y'_i, Y''_i)$ with $Y'_i := -\theta(X'_i), Y''_i := -\theta(X''_i)$. [19] It is enough to check that $\mathfrak{g}$ is the graph of a θ-equivariant isomorphism of Lie algebras; note that if such a morphism exists, then it is uniquely defined and restricts on $\mathfrak{t}'_{\mathbb{C}}$ to $d\kappa$. The θ-equivariance will follow from the construction of $\mathfrak{g}$. By applying the same argument with the roles of $\mathfrak{g}'$ and $\mathfrak{g}''$ reversed, we reduce to showing that $\mathfrak{g}$ is the graph of a morphism of Lie algebras. Let $\pi' : \mathfrak{g} \to \mathfrak{g}'$ and $\pi'' : \mathfrak{g} \to \mathfrak{g}''$ denote the projections. By Corollary 145, we see that π' and π'' are surjective. We might thus hope to define a morphism $f : \mathfrak{g}' \to \mathfrak{g}''$ by requiring that $f(x) := y$ whenever $(x, y) \in \mathfrak{g}$. Since $\mathfrak{g}$ is a subalgebra, we know that f is a morphism provided that it is well-defined as a set-theoretic map. We thus need only check that if (x, y_1) and (x, y_2) both belong to $\mathfrak{g}$, then $y_1 = y_2$. Equivalently, we need to check that if $(0, y)$ belongs to $\mathfrak{g}$, then $y = 0$; in other words, we must show that $\ker(\pi'')$ is trivial.

We claim that

$$\mathfrak{g} \cap (\mathfrak{t}'_{\mathbb{C}} \oplus \mathfrak{t}''_{\mathbb{C}}) = \mathfrak{t}_{\mathbb{C}}. \tag{80}$$

Before proving the claim, we explain why it suffices. Suppose for the sake of contradiction that $\ker(\pi'')$ is nonzero. Then $\mathfrak{a}' := \pi'(\ker(\pi''))$ is a nonzero ideal of $\mathfrak{g}'$. By Corollary 146, we have $\mathfrak{a}' \cap \mathfrak{t}'_{\mathbb{C}} \neq \{0\}$. We may thus find $0 \neq x' \in \mathfrak{t}'_{\mathbb{C}}$ so that $(x', 0) \in \ker(\pi'')$. Clearly

$$(x', 0) \in \mathfrak{g} \cap (\mathfrak{t}'_{\mathbb{C}} \oplus \mathfrak{t}''_{\mathbb{C}}).$$

From the claim (80), it follows that $(x', 0) \in \mathfrak{t}_{\mathbb{C}}$, whence that $0 = d\kappa(x')$. Since $d\kappa$ is an isomorphism, this forces $x' = 0$, giving the required contradiction.

We turn finally to the proof of the claim (80). Let $\mathfrak{h} \subseteq \mathfrak{g}' \oplus \mathfrak{g}''$ denote the set defined as on the RHS of (77), with e_i, f_i defined in terms of $X_i, Y_i \in \mathfrak{g}$ as in (78), (79). Then $\mathfrak{t}_{\mathbb{C}} \subseteq \mathfrak{h}$; moreover, $\mathfrak{t}_{\mathbb{C}}$ is self-centralizing in $\mathfrak{h}$, since the $\mathfrak{t}_{\mathbb{C}}$-weights of $\oplus \mathbb{C}e_i$ (resp. of $\oplus \mathbb{C}f_i$) are positive (resp. negative). In particular, $(\mathfrak{t}'_{\mathbb{C}} \oplus \mathfrak{t}''_{\mathbb{C}}) \cap \mathfrak{h} = \mathfrak{t}_{\mathbb{C}}$, so to establish (80), we reduce to showing that $\mathfrak{g} = \mathfrak{h}$. To that end, observe first that $\mathfrak{h}$ is contained in $\mathfrak{g}$ and contains the generating set $S := \mathfrak{t}_{\mathbb{C}} \cup \{X_i\} \cup \{Y_i\}$ for $\mathfrak{g}$. We thereby reduce to verifying that $\mathfrak{h}$ is a subalgebra. Since $\mathfrak{h} \subseteq \mathfrak{g}$, we may reduce further to showing that $\mathfrak{h}$ is closed under ad_Z for each $Z \in S$. For this we appeal to the relations noted in Corollary 145, applied to each component $\mathfrak{g}', \mathfrak{g}''$ of $\mathfrak{g}' \oplus \mathfrak{g}''$, together with the assumed compatibility between γ'_i and γ''_i.

---

[19] End of lecture #18, Tuesday, 30 Apr 2019

This completes the proof.

7.21. **Dynkin diagrams.** They provide a convenient pictorial representation of the Cartan matrix. Here we just summarize what was discussed in class; look at Wikipedia or any of the course references for a detailed discussion.

The vertices of the Dynkin diagram correspond to the simple roots. For any pair (α, β) of *distinct* simple roots, we have seen that there are at most seven possibilities for the pair $(N_{\alpha\beta}, N_{\beta\alpha})$ of integers appearing in the Cartan matrix:

- $(0,0)$: we draw no edge between α and β
- $(-1,-1)$: we draw a simple undirected edge between α and β
- $(-1,-2)$: we draw a double oriented edge from α to β
- $(-2,-1)$: we draw a double oriented edge from β to α
- $(-1,-3)$: we draw a triple oriented edge from α to β
- $(-3,-1)$: we draw a triple oriented edge from β to α

We illustrated with $K = \mathrm{U}(n)$ or $\mathrm{SU}(n)$ or (n), where the Cartan matrix looks like (for $n = 5$)

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix},$$

so the Dynkin diagram looks like a series of vertices connected by simple edges.

We've focused in this class on Lie groups rather than Lie algebras. A variant of Theorem 141 (with similar proof) is that if K', K'' are connected compact Lie groups with maximal tori T', T'' of the same dimension, then $\mathfrak{k}' \cong \mathfrak{k}''$ iff the Cartan matrices are the same (up to relabeling indices) iff the Dynkin diagrams are the same.

7.22. **Root data.**

Definition 148. A *root datum* is a quadruple $\Psi = (X, \Phi, X^\vee, \Phi^\vee)$ consisting of

(i) finitely-generated free abelian groups $X, X^\vee$ of the same rank n, thus $X \cong \mathbb{Z}^n, X^\vee \cong \mathbb{Z}^n$,

(ii) a perfect $\mathbb{Z}$-linear pairing $X \otimes_{\mathbb{Z}} X^\vee \xrightarrow{(,)} \mathbb{Z}$,

(iii) finite subsets $\Phi \subseteq X, \Phi^\vee \subseteq X$, and

(iv) a bijection $\Phi \leftrightarrow \Phi^\vee$ denoted $\alpha \leftrightarrow \alpha^\vee$ such that

(v) $(\alpha, \alpha^\vee) = 2$ for all $\alpha \in \Phi$,

(vi) $s_\alpha(\lambda) := \lambda - (\lambda, \alpha^\vee)\alpha$ defines a linear automorphism $s_\alpha : X \to X$ such that $s_\alpha(\Phi) = \Phi$, and

(vii) $s_{\alpha^\vee}(\lambda^\vee) := \lambda^\vee - (\alpha, \lambda^\vee)\alpha^\vee$ defines a linear automorphism $s_{\alpha^\vee} : X^\vee \to X^\vee$ such that $s_{\alpha^\vee}(\Phi^\vee) = \Phi^\vee$.

We say that Ψ is *reduced* if $\mathbb{Q}\alpha \cap \Phi = \{\alpha, -\alpha\}$ for all $\alpha \in \Phi$.

There is an obvious notion of an isomorphism of root data.

For example, if K is a compact connected Lie group with maximal torus T, then

$$\Psi(K : T) := (X(T), \Phi(K : T), X^\vee(T), \Phi^\vee(K : T))$$

is a root datum; here $\Phi^\vee(K : T) := \{\alpha^\vee : \alpha \in \Phi(K : T)\}$. Using the conjugacy of maximal tori, we see that the isomorphism class of $\Psi(K : T)$ is independent of T.

Theorem 149. *The map*

$$\{compact\ connected\ Lie\ groups\ K\}/\sim\ \longrightarrow\ \{reduced\ root\ data\ \Psi\}/\sim,$$

given by sending K to $\Psi(K:T)$ for some maximal torus T, is a bijection.

Sketch of proof. Injectivity follows readily from Theorem 141 (exercise). We omit the proof of surjectivity; one way to proceed would be to argue as in the proof of Theorem 141, but incorporating the Serre relations $\mathrm{ad}_{X_i}^{-N_{ij}+1} X_j = \mathrm{ad}_{Y_i}^{-N_{ij}+1} Y_j = 0$. $\qquad\square$

[20]

8. Representations of compact Lie groups

Let K be a compact connected Lie group. Our aim in this section is to generalize the results established in §3 when $K = \mathrm{U}(n)$.

8.1. **Setup and preliminaries.** Choose a maximal torus T and then a Weyl chamber $C \subseteq \mathfrak{t}_{\mathbb{R}}^{\mathrm{reg}}$. These choices define a Weyl group W and sets $\Phi \supseteq \Phi^+ \supseteq \Delta$ of roots, positive roots and simple roots, respectively. Recall that C defines a notion of $\lambda \in \mathfrak{t}_{\mathbb{R}}^*$ being *dominant*, which means that it belongs to the closure $\overline{C^\vee} = \{\lambda : \lambda(H_\alpha) \geq 0 \text{ for all } \alpha \in \Phi^+\}$ of the Weyl chamber $C^\vee = \{\lambda : \lambda(H_\alpha) > 0 \text{ for all } \alpha \in \Phi^+\}$, and *strictly dominant* if in fact $\lambda \in C^\vee$. Recall that C defines a partial on $\mathfrak{t}_{\mathbb{R}^*}$, where we say that $\lambda_1 \geq \lambda_2$ if $(\lambda_1 - \lambda_2)(C) \geq 0$, or equivalently, if $\lambda_1 - \lambda_2 \in \sum_{\alpha \in \Phi^+} \mathbb{R}_{\geq 0}\alpha$; we say also that λ is *positive* if $\lambda \geq 0$.

We remark that one can characterize the subsets $\Phi^+ \subseteq \Phi$ arising in the above way as those for which

- $\alpha, \beta \in \Phi^+, \alpha + \beta \in \Phi \implies \alpha + \beta \in \Phi^+$ and
- Φ is the disjoint union of Φ^+ and $-\Phi^+$.

(See any of the main course references.) The choice of Φ^+ determines $C = \{x : \alpha(x) > 0 \text{ for all } \alpha \in \Phi^+\}$ and $C^\vee$. A convenient way to choose Φ^+ is via a *lexicographical ordering*: choose a faithful embedding $K \hookrightarrow \mathrm{U}(n)$, hence an embedding $\mathfrak{t}_{\mathbb{R}} \hookrightarrow \mathbb{R}^n$, identify $\mathfrak{t}_{\mathbb{R}}^*$ with $\mathfrak{t}_{\mathbb{R}}$ in the usual way via the pairing $(x, y) \mapsto \mathrm{trace}(xy)$ (the restriction of the standard Euclidean norm on $\mathbb{R}^n$), and equip $\mathfrak{t}_{\mathbb{R}}^*$ with the lexicographical order coming from $\mathbb{R}^n$. This is what we did earlier (implicitly) in the case $K = \mathrm{U}(n)$.

We recall some additional facts established in the preceeding section:

(i) W acts simply-transitively on the set of Weyl chambers.
(ii) (dual form of Corollary 130) Any $\lambda \in \mathfrak{t}_{\mathbb{R}}^*$ is W-conjugate to a unique dominant element.
(iii) Any dominant $\lambda \in \mathfrak{t}_{\mathbb{R}}^*$ satisfies $\lambda \geq w(\lambda)$ for all $w \in W$.

Recall finally the Weyl integral formula (§7.6)

$$\int_K f = \frac{1}{|W|} \int_{g \in K/T} \int_{t \in T} D(t) f(gtg^{-1}), \tag{81}$$

where we integrate with respect to probability Haar measures and where the Jacobian factor D is given explicitly by

$$D(t) = |\det(t - 1|\mathfrak{g}/\mathfrak{t}_{\mathbb{C}})| = \prod_{\alpha \in \Phi} |t^\alpha - 1|.$$

---

[20]End of lecture #???

Our first aim is to factor $D = \Delta \overline{\Delta} = |\Delta|^2$ as in the case $K = \mathrm{U}(n)$. Such a factorization holds formally with

$$\Delta(t) := \prod_{\alpha > 0} (t^{\alpha/2} - t^{-\alpha/2}).$$

Unfortunately, Δ is not in general well-defined as a function on T; the source of this issue is that the expressions $t^{\alpha/2}$ themselves are not well-defined, since it may happen that $\alpha \in \mathfrak{t}_{\mathbb{Z}}^*$ but $(1/2)\alpha \notin \mathfrak{t}_{\mathbb{Z}}^*$. To circumvent this issue, we pass to a covering torus $\tilde{T}$ of T. Recall that the normalized exponential map $e(x) := \exp(2\pi i x)$ induces an isomorphism $e : \mathfrak{t}_{\mathbb{R}}/\mathfrak{t}_{\mathbb{Z}} \xrightarrow{\sim} T$. We denote by $\tilde{T}$ the covering torus of T corresponding to the sublattice $2\mathfrak{t}_{\mathbb{Z}}$ of $\mathfrak{t}_{\mathbb{Z}}$; it comes with an isomorphism $e : \mathfrak{t}_{\mathbb{R}}/2\mathfrak{t}_{\mathbb{Z}} \xrightarrow{\sim} \tilde{T}$ and a canonical surjection $\tilde{T} \to T$. In other words, choosing a basis for $\mathfrak{t}_{\mathbb{Z}}$, we may identify T with $\mathbb{R}^n/\mathbb{Z}^n$ and $\tilde{T}$ with $\mathbb{R}^n/2\mathbb{Z}^n$. In any event, for $\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^*$, we have $\lambda(2\mathfrak{t}_{\mathbb{Z}}) \subseteq \mathbb{Z}$, and so the character $e^\lambda : \tilde{T} \to \mathrm{U}(1)$ given by $\exp(x) \mapsto \exp(\lambda(x))$ is well-defined. In particular, the above formula defines $\Delta : \tilde{T} \to \mathbb{C}$. We identify $D : T \to \mathbb{C}$ with a function $D : \tilde{T} \to \mathbb{C}$ via pullback. The identity $|\Delta|^2 = D$ then holds. For a class function f on K, the Weyl integral formula (81) implies that

$$\langle f, f \rangle_K = \frac{1}{W} \langle \Delta f, \Delta f \rangle_{\tilde{T}},$$

where we identify f with a function on $\tilde{T}$ by restricting to T and then pulling back to $\tilde{T}$, and where we integrate with respect to the probability Haar on $\tilde{T}$.

Set

$$\rho := \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha.$$

It defines an element of $(1/2)\mathfrak{t}_{\mathbb{Z}}^*$, hence a character of $\tilde{T}$. We may rewrite

$$\Delta(t) = t^\rho \prod_{\alpha > 0} (1 - t^{-\alpha})$$

We pause to record some basic properties of ρ.

Lemma 150. *(i) ρ is strictly dominant. Its W-stabilizer is trivial.*
(ii) For all $w \in W$, we have

$$\rho - w(\rho) = \sum_{\substack{\alpha \in \Phi: \\ \alpha > 0, \\ w^{-1}(\alpha) < 0}} \alpha. \tag{82}$$

In particular, $\rho - w(\rho)$ defines a positive element of $\mathfrak{t}_{\mathbb{Z}}^*$.

Proof. We first establish the formula (82). We have

$$w(\rho) = \frac{1}{2} \sum_{\alpha > 0} w(\alpha) = \sum_{\alpha : w^{-1}(\alpha) > 0} \alpha = \sum_{\alpha > 0} \varepsilon_\alpha \alpha,$$

where

$$\varepsilon_\alpha = \begin{cases} 1 & \text{if } w^{-1}(\alpha) > 0, \\ -1 & \text{if } w^{-1}(\alpha) < 0. \end{cases}$$

Thus
$$\rho - w(\rho) = \sum_{\alpha > 0} \frac{1 - \varepsilon_\alpha}{2} \alpha.$$

We have
$$\frac{1 - \varepsilon_\alpha}{2} = \begin{cases} 1 & \text{if } w^{-1}(\alpha) < 0, \\ 0 & \text{otherwise,} \end{cases}$$

which leads to the required formula. It follows immediately that $\rho - w(\rho)$ defines a positive element of $\mathfrak{t}_{\mathbb{Z}}^*$.

Suppose now that $w \in W$ fixes ρ. Then w stabilizes Φ^+, hence w stabilizes C. Since we have seen (in Lemma 116) that W acts freely on the set of Weyl chambers, it follows that $w = 1$. Thus ρ has trivial W-stabilizer.

It remains only to verify that ρ is strictly dominant, i.e., that for each $\alpha \in \Phi^+$, we have $\rho(H_\alpha) > 0$. To that end, note first that since ρ has trivial W-stabilizer, it is not fixed by the root reflection s_α, and so $\rho(H_\alpha) \neq 0$. Suppose for the sake of contradiction that $\rho(H_\alpha) < 0$. Then $s_\alpha(\rho) - \rho = -\rho(H_\alpha)\alpha > 0$. But we have seen already using the formula (82) that $s_\alpha(\rho) - \rho < 0$, giving the required contradiction. $\qquad\square$

Let L denote the ring of Laurent polynomials on $\tilde{T}$, thus
$$L = \oplus_{\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^*} \mathbb{C}e^\lambda,$$

where as usual $e^\lambda = [t \mapsto t^\lambda]$. The Weyl group W acts on L. We set
$$L^{\mathrm{sym}} := \{f \in L : w(f) = f \text{ for all } w \in W\}.$$

For $w \in W$, we denote by $(-1)^w$ the determinant of the action of w on $\mathfrak{t}_{\mathbb{R}}$. For instance, for the root reflections s_α, we have $(-1)^{s_\alpha} = -1$, because reflections have determinant -1. More generally, if $w \in W$ may be written as a product of exactly $k \geq 0$ root reflections, then $(-1)^w = (-1)^k$. This observation justifies the notation. We set
$$L^{\mathrm{alt}} := \{f \in L : w(f) = (-1)^w f \text{ for all } w \in W\}.$$

Since the (simple) root reflections generate W, we may also write
$$L^{\mathrm{alt}} = \{f \in L : s_\alpha(f) = -f \text{ for all } \alpha \in \Delta\}.$$

For $\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^*$, define
$$A(\lambda) := \sum_{w \in W} (-1)^w e^{w(\lambda)}.$$

Lemma 151.   *(i)* $A(\lambda) \in L^{\mathrm{alt}}$ *for all* $\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^*$.
(ii) $L^{\mathrm{alt}} = \oplus_{\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^* \cap C^\vee} \mathbb{C}A(\lambda)$.

Proof.   (i) Direct calculation.
   (ii) Let $f \in L^{\mathrm{alt}}$. Then $f = \sum_{\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^*} c(\lambda)e^\lambda$, where $c(w(\lambda)) = (-1)^w c(\lambda)$ for all $w \in W$. If λ is irregular, i.e., if $\lambda(H_\alpha) = 0$ for some root α, then $s_\alpha(\lambda) = \lambda$, while $(-1)^{s_\alpha} = -1$, and so $c(\lambda) = 0$. If λ is regular, then its unique dominant W-conjugate $w(\lambda)$ belongs to $C^\vee$, and we have $c(\lambda) = (-1)^w c(w(\lambda))$. It follows that $f = \sum_{\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^* \cap C^\vee} c(\lambda)A(\lambda)$. The uniqueness of this decomposition follows from the fact that if $\lambda \in (1/2)\mathfrak{t}_{\mathbb{Z}}^* \cap C^\vee$, then $w(\lambda) \notin C^\vee$ for all nontrivial $w \in W$.
$\qquad\square$

Let π be a finite-dimensional representation of G. Recall that the *weights* of π are those $\mu \in X(T)$ for which the weight space $\pi[\mu] := \{v \in \pi : tv = t^\mu v \text{ for all } t \in T\}$ is nonzero. The character χ_π of π is a class function on G, which we may identify with a W-invariant function on T, and in fact an element $\chi_\pi \in L^{\mathrm{sym}}$, given by $\chi_\pi = \sum_{\mu \in X(T)} m_\pi(\mu) e^\mu$. For any element $f \in L$, we may speak more generally of the *weights* of f (i.e., those $\mu \in X(\tilde{T})$ for which the coefficient of e^μ in f is nonzero), and we say that λ is

- a *maximal weight* of f if there does not exist a weight μ of f with $\mu > \lambda$, and
- the *highest weight* of f if every weight μ of f satisfies $\mu \leq \lambda$.

Note that f has (in general, many) maximal weights, and that f has a highest weight iff it has a unique maximal weight. This terminology applies also to π in place of f by considering the character χ_π. For instance, $\Delta = \prod_{\alpha>0}(e^{\alpha/2} - e^{-\alpha/2}) \in L^{\mathrm{alt}}$ and $A(\rho) = \sum_{w \in W}(-1)^w e^{w(\rho)} = A(\rho)$ are each of the form $e^\rho + \cdots$ where $\cdots$ denotes the contributions of weights $w(\rho)$ strictly less than ρ, and so Δ and $A(\rho)$ both have highest weight ρ.

Lemma 152.   (i) $\Delta \in L^{\mathrm{alt}}$.
(ii) For all $f \in L^{\mathrm{alt}}$, we have $f/\Delta \in L^{\mathrm{sym}}$.
(iii) $\Delta = A(\rho)$.
(iv) Let λ be a dominant element of $(1/2)\mathfrak{t}_{\mathbb{Z}}^*$. Then $\lambda + \rho$ is a strictly dominant element $(1/2)\mathfrak{t}_{\mathbb{Z}}^*$, and

$$\frac{A(\lambda+\rho)}{A(\rho)}(1) = \prod_{\alpha>0} \frac{\langle \lambda+\rho, H_\alpha \rangle}{\langle \rho, H_\alpha \rangle}. \tag{83}$$

Proof.   (i) It is enough to check for each simple root β that

$$\Delta(s_\beta(t)) = -\Delta(t).$$

Using that $s_\beta^2 = 1$, we compute that

$$\Delta(s_\beta(t)) = \prod_{\alpha>0}(t^{s_\beta(\alpha)/2} - t^{-s_\beta(\alpha)/2}).$$

Since $s_\beta(\beta) = -\beta$, we have

$$(t^{s_\beta(\beta)/2} - t^{-s_\beta(\beta)/2}) = -(t^{\beta/2} - t^{-\beta/2}).$$

Let $\alpha \in \Phi^+ - \{\beta\}$. Then (see the homework for details concerning the following terminology and facts) the Weyl chambers C and $s_\beta(C)$ are adjacent, separated by the wall $\beta^\perp$ and no other wall. In particular, α takes the same sign on C and $s_\beta(C)$, hence α is positive on $s_\beta(C)$; equivalently, $s_\beta(\alpha)$ is positive on C, i.e., $s_\beta(\alpha) > 0$. Thus s_β acts on $\Phi^+ - \{\beta\}$, and so

$$\prod_{\alpha>0, \alpha\neq\beta}(t^{s_\beta(\alpha)/2} - t^{-s_\beta(\alpha)/2}) = \prod_{\alpha>0, \alpha\neq\beta}(t^{\alpha/2} - t^{-\alpha/2}).$$

The required identity follows.

(ii) It is enough to show that f is divisible by Δ in L. We record two proofs (the second of which was given in lecture, the first of which I find a bit more natural). Both proofs use the fact that for distinct positive roots α and β,

$$\mathbb{C}\alpha \cap \mathbb{C}\beta = \{0\}. \tag{84}$$

(To see this, it suffices to check that $\beta \notin \mathbb{C}\alpha$, which follows from the identity $\mathbb{C}\alpha \cap \Phi = \{\alpha, -\alpha\}$ and our assumptions.)

(a) (First proof) Set $Z := \{z \in (\mathbb{C}^\times)^n : \Delta(z) = 0\}$. Then $z \in Z$ iff $z^\alpha = 1$ for some positive root α. The identity $z^\alpha = 1$ is equivalent to $s_\alpha(z) = z$. Let $f \in L^{\mathrm{alt}}$ and $z \in Z$. Then, choosing α so that $s_\alpha(z) = z$, we have $f(z) = f(s_\alpha(z)) = -f(z)$, hence $f(z) = 0$. Thus f vanishes on Z. By the Nullstellensatz, it follows that f belongs to the radical of the ideal in Δ generated by Δ. Note that L is a UFD (being a localization of the UFD $\mathbb{C}[t_1, \ldots, t_n]$). To conclude, it suffices to show that Δ defines a squarefree element of L, i.e., that no prime factor of Δ occurs with multiplicity > 1. We first compute the prime factorization of $t^\alpha - 1$ for each nontrivial character α of $\tilde{T}$. The character group $X(\tilde{T})$ is isomorphic to $\mathbb{Z}^n$, where $n = \dim(\tilde{T})$. Suppose that α is of the form $d\beta$ for some natural number d and some character β of $\tilde{T}$, with d chosen maximally. Then

$$t^\alpha - 1 = \prod(t^\beta - \zeta), \tag{85}$$

with the product taken over all dth roots of unity ζ. We claim that each factor $t^\beta - \zeta$ is irreducible. For this it suffices to show more generally that if α is a *primitive* nontrivial character of $\tilde{T}$ (thus α corresponds to a nonzero element of $\mathbb{Z}^n$ whose coordinates are relatively prime) and $\zeta \in \mathbb{C}^\times$, then $t^\alpha - \zeta$ is irreducible. Indeed, since α is primitive, we may apply a linear change of variables on $\mathbb{Z}^n$ to reduce to the case that $\alpha = (1, 0, \ldots, 0)$, so that $t^\alpha - \zeta = t_1 - \zeta$, which is clearly irreducible. (The existence of such a change of variables is a simple exercise in linear algebra, closely related to the fact that finitely-generated torsion-free modules over $\mathbb{Z}$ are free; indeed, the latter fact implies that the quotient $X(\tilde{T})/\mathbb{Z}\alpha$ is free, hence that the surjection to that quotient from $X(\tilde{T})$ splits, leading to a direct sum decomposition $X(\tilde{T}) = \mathbb{Z}\alpha \oplus (\cdots)$ and hence the required change of variables.) Since no prime factor in (85) occurs twice, we see in particular that $t^\alpha - 1$ is squarefree. To establish the same for Δ, we need to check that no prime factor of $t^\alpha - 1$ divides $t^\beta - 1$ whenever α and β are distinct positive roots. This follows from (85) and (84).

(b) (Second proof) We show first that if $f \in L$ vanishes on $\{z : z^\alpha = 1\}$, then $f/(t^\alpha - 1) \in L$. Choosing coordinates $t_1, \ldots, t_n$ on $\tilde{T}$, we may represent α as $(\alpha_1, \ldots, \alpha_n)$ for some integers α_j. We may assume for convenience, by inverting the coordinate t_n if necessary, that $\alpha_n \geq 0$. Setting $R := \mathbb{Z}[t_1^{\pm 1}, \ldots, t_{n-1}^{\pm 1}]$, we then have $t^\alpha - 1 \in R[t_n]$. We may assume, after multiplying f by a sufficiently large power of t_n, then $f \in R[t_n]$. By division with remainder in the polynomial t_n, we then have $f = (t_n^{\alpha_n} - \prod_{j<n} t_j^{-\alpha_j})q + r$ for some $q, r \in R[t_n]$ for which the degree in t_n of r is strictly less than α_n. This identity implies that r vanishes on the solution set to the equation $t^\alpha = 1$. For any given values $t_1, \ldots, t_{n-1} \in \mathbb{C}^\times$ of the first $n - 1$ coordinates, that equation has exactly α_n solutions in the variable t_n. The noted degree bound on r thus implies that $r = 0$. It follows as required that f is divisible by $t^\alpha - 1$. From (84) we see that for distinct positive roots α and β, the tangent planes to the hypersurfaces $\{t : t^\alpha = 1\}$ and $\{t : t^\beta = 1\}$ are traverse at

any point of the intersection $\{t : t^\alpha = t^\beta = 1\}$. Thus if f vanishes on $\{z : z^\beta = 1\}$ and is divisible by $t^\alpha - 1$, then $f/(t^\alpha - 1)$ likewise vanishes on $\{z : z^\beta = 1\}$. We may thus iteratively apply the argument of the preceeding paragraph to obtain the required conclusion.

(iii) By the previous item, we have $A(\rho)/\Delta \in L^{\mathrm{sym}}$. We've noted already that Δ and $A(\rho)$ each have ρ as highest weight. If λ is any maximal weight of $A(\rho)/\Delta$, then $\lambda + \rho = \rho$, whence $\lambda = 0$. Thus $A(\rho)/\Delta$ is a complex scalar; since $A(\rho)$ and Δ are each of the form $e^\rho + (\cdots)$, we conclude that $A(\rho)/\Delta = 1$.

(iv) We have seen that ρ is strictly dominant. Since λ is dominant, it follows that $\lambda + \rho$ is strictly dominant. It remains to check the formula (83). This requires taking a limit as in the proof of the analogous result for $\mathrm{U}(n)$ (Lemma 42). Let us first rewrite the identity $\Delta = A(\rho)$ established above in the form: for $x \in \mathfrak{t}_\mathbb{C}$,

$$\sum_W (-1)^w e^{\langle w(x),\rho \rangle} = \prod_{\alpha>0} (e^{\alpha(x)/2} - e^{-\alpha(x)/2}). \tag{86}$$

A similar argument (applied with the roles of $\mathfrak{t}_\mathbb{C}$ and $\mathfrak{t}_\mathbb{C}^*$ reversed) gives for $\lambda \in \mathfrak{t}_\mathbb{C}^*$ that

$$\sum_W (-1)^w e^{\langle w(\lambda),H_\rho \rangle} = \prod_{\alpha>0} (e^{\lambda(H_\alpha)/2} - e^{-\lambda(H_\alpha)/2}). \tag{87}$$

It follows that for $t = \exp(\varepsilon H_\rho)$ with $\varepsilon > 0$ small, we have

$$\sum_W (-1)^w t^{w(\lambda)} = \prod_{\alpha>0} (e^{\varepsilon\lambda(H_\alpha)/2} - e^{-\varepsilon\lambda(H_\alpha)/2}) \sim \varepsilon^{|\Phi_+|} \prod_{\alpha>0} \langle \lambda, H_\alpha \rangle.$$

Thus

$$\frac{\sum_W (-1)^w t^{w(\lambda+\rho)}}{\sum_W (-1)^w t^{w(\rho)}} \Big|_{t=1} = \prod_{\alpha>0} \frac{\langle \lambda + \rho, H_\alpha \rangle}{\langle \rho, H_\alpha \rangle},$$

as required.

$\square$

8.2. Classification of irreducibles.

Theorem 153. *Let K be a compact connected Lie group with maximal torus T, hence root set Φ and Weyl group W. Choose a Weyl chamber C, hence sets Φ^+ and Δ of positive and simple roots, respectively, as well as a partial order $\geq$ on $\mathfrak{t}_\mathbb{R}^*$.*

Each $\pi \in \mathrm{Irr}(K)$ has a highest weight $\lambda \in X(T) = \mathfrak{t}_\mathbb{Z}^*$; this is a dominant element for which the Weyl character formula

$$\chi_\pi = \frac{A_{\lambda+\rho}}{A_\rho}$$

holds, with

$$A_\mu := \sum_{w \in W} (-1)^w e^{w(\mu)} \in X(\tilde{T}) = \frac{1}{2}\mathfrak{t}_\mathbb{Z}^*, \quad \rho := \frac{1}{2}\sum_{\alpha>0} \alpha.$$

The map $\pi \mapsto \lambda$ defines a bijection

$$\mathrm{Irr}(K) \leftrightarrow \{dominant\ \lambda \in X(T)\} \leftrightarrow X(T)/W.$$

Proof. Let $\pi \in \mathrm{Irr}(K)$. Let $\lambda \in X(T)$ be any maximal weight of π. Since χ_π is W-invariant, and since every element μ of $X(T)$ is W-conjugate to a unique dominant element which is then $\geq \mu$, we know that λ is dominant. Since Δ has highest weight

ρ with $\Delta = e^\rho + \cdots$, we know that $\Delta \chi_\pi \in L^{\mathrm{alt}}$ has $\lambda + \rho$ as a maximal weight, with $\Delta \chi_\pi = m_\pi(\lambda) e^{\lambda + \rho} + \cdots$. Regrouping terms into monomial alternating functions gives

$$\Delta \chi_\pi = m_\pi(\lambda) A_{\lambda + \rho} + \cdots, \tag{88}$$

where $\cdots$ denotes a linear combination of $A_{\mu + \rho}$ taken over μ with $\mu + \rho$ strictly dominant and μ not greater than λ. By the Weyl integral formula and Schur orthogonality, we have (as in the case of $\mathrm{U}(n)$)

$$1 = \langle \chi_\pi, \chi_\pi \rangle_G = \frac{\langle D\chi_\pi, \chi_\pi \rangle_T}{|W|} = \frac{\langle \Delta \chi_\pi, \Delta \chi_\pi \rangle_{\tilde{T}}}{|W|} = \frac{\|m_\pi(\lambda) A_{\lambda + \rho} + \cdots\|^2}{|W|} \geq |m_\pi(\lambda)|^2. \tag{89}$$

Here we used that $\lambda + \rho$ is strictly dominant and that for strictly dominant λ_1, λ_2,

$$\langle A_{\lambda_1}, A_{\lambda_2} \rangle = \begin{cases} |W| & \text{if } \lambda_1 = \lambda_2, \\ 0 & \text{otherwise} \end{cases}.$$

Since $m_\pi(\lambda) \in \mathbb{Z}_{\geq 1}$, we deduce from (89) that $m_\pi(\lambda) = 1$ and hence that equality holds in each step, i.e., that the $\cdots$ in (88) vanishes. This gives the required formula for χ_π in terms of λ. The proof shows also that λ is uniquely determined, hence is the unique maximal weight, i.e., the highest weight of π. We thus obtain an injective map $\mathrm{Irr}(K) \to \{\text{dominant } \lambda \in X(T)\}$ as in the statement of the theorem; to complete the proof, it remains only to show that this map is surjective. We imitate the first proof given earlier for $\mathrm{U}(n)$. Suppose λ is a dominant element of $X(T)$ not arising as the highest weight of any element of $\mathrm{Irr}(K)$. By Lemma 152, the ratio $A_{\lambda + \rho}/A_\rho$ defines an element of L^{sym}, hence in particular a function $\tilde{T} \to \mathbb{C}$. We claim that this function factors through T. To see this, we write

$$\frac{A_{\lambda + \rho}}{A_\rho} = \frac{\sum_{w \in W} (-1)^w e^{w(\lambda + \rho) - \rho}}{e^\rho \prod_{\alpha > 0} (1 - e^{-\alpha})},$$

and observe that having rewritten the ratio in this way, both the numerator and denominator factor through T: for the denominator, this is clear, while for the numerator, we use that $w(\lambda + \rho) - \rho = w(\lambda) - (\rho - w(\rho))$, which belongs to $\mathfrak{t}_\mathbb{Z}^*$ thanks to Lemma 150. This completes the verification of the claim. Thus $A_{\lambda + \rho}/A_\rho$ defines a W-invariant function $T \to \mathbb{C}$, hence a class function $K \to \mathbb{C}$. For each $\pi \in \mathrm{Irr}(K)$, say with highest weight μ, then the Weyl integral formula gives

$$\langle A_{\lambda + \rho}/A_\rho, \chi_\pi \rangle_K = \langle A_{\lambda + \rho}, \Delta \chi_\pi \rangle_{\tilde{T}} = \langle A_{\lambda + \rho}, A_{\mu + \rho} \rangle_{\tilde{T}} = 0,$$

using in the last step that λ, μ are dominant and distinct. The Peter–Weyl theorem then implies that $A_{\lambda + \rho}/A_\rho = 0$, which is absurd. This completes the proof of the required contradiction. $\square$

Remark 154. We indicate a way to formulate the conclusion of Theorem 153 without explicit reference to a maximal torus. Recall from §7.22 that there is a bijection

$$\{\text{compact connected } K\}/\sim \quad \leftrightarrow \quad \{\text{reduced root data } \Psi = (X, \Phi, X^\vee, \Phi^\vee)\}$$

$$K \mapsto \Psi(K) := (X(T), \Phi(K : T), X^\vee(T), \Phi^\vee).$$

There is a "duality" $\Psi \mapsto \Psi^\vee$ on the set of (reduced) root data given by

$$\Psi^\vee := (X^\vee, \Phi^\vee, X, \Phi).$$

We may define the dual group $K^\vee$ of K to be the connected compact Lie group (up to isomorphism) having the dual root datum, i.e., $\Psi(K^\vee) = \Psi(K)^\vee$; it comes with a maximal torus $T^\vee \subseteq K^\vee$ having natural identifications $X(T^\vee) = X^\vee(T)$ and $X^\vee(T^\vee) = X(T)$, and we may furthermore identify the Weyl group W of T with the Weyl group of $T^\vee$. We then have natural bijections

$$\begin{aligned}
\mathrm{Irr}(K) &\cong \{\text{dominant } \lambda \in X^\vee(T)\} \\
&\cong X(T)/W \\
&\cong X^\vee(T^\vee)/W \\
&\cong \mathrm{Hom}(\mathrm{U}(1), K^\vee)/\sim,
\end{aligned}$$

where the equivalence relation $\sim$ in the last expression is given by conjugation.

8.3. **Borel–Weil.** Recall the usual notation: K is a compact connected Lie group with Lie algebra $\mathfrak{k}$, complexification G, and complexified Lie algebra $\mathfrak{g}$. Recall that $\mathrm{Irr}(K)$ is in natural bijection with the set of isomorphism classes of holomorphic (equivalently, algebraic) finite-dimensional irreducible representation of G. We fix a maximal torus $T \leq K$, with Lie algebra denoted $\mathfrak{t}$, and set $\mathfrak{t}_\mathbb{R} := i\mathfrak{t} \subseteq \mathfrak{t}_\mathbb{C}$, as usual. We denote by $A \leq G$ the connected (complex) Lie subgroup with Lie algebra $\mathrm{Lie}(A) = \mathfrak{t}_\mathbb{C}$. For example, if $K = \mathrm{U}(n)$, so that $G = \mathrm{GL}_n(\mathbb{C})$, then with the choice

$$T = \begin{pmatrix} \mathrm{U}(1) & & \\ & \ddots & \\ & & \mathrm{U}(1) \end{pmatrix}$$

we have

$$A = \begin{pmatrix} \mathbb{C}^\times & & \\ & \ddots & \\ & & \mathbb{C}^\times \end{pmatrix}.$$

Fix a Weyl chamber $C \subseteq \mathfrak{t}_\mathbb{R}$ and hence choices of positive roots Φ^+ and simple roots Δ, as usual. Set

$$\mathfrak{n} := \oplus_{\alpha \in \Phi^+}$$

and

$$\mathfrak{b} := \mathfrak{n} \oplus \mathfrak{t}_\mathbb{C} \leq \mathfrak{g}.$$

(The directness of the sum follows immediately from the root space decomposition of $\mathfrak{g}$.) For example, if we take $K = \mathrm{U}(n)$, T as above, and for C the "standard" Weyl chamber $\{(x_1, \ldots, x_n) : x_i > x_j \text{ for } i < j\}$, then $\Phi^+ = \{\varepsilon_i - \varepsilon_j : i < j\}$, and (for $n = 3$, say)

$$\mathfrak{n} = \oplus_{i<j}\mathbb{C}E_{ij} = \begin{pmatrix} 0 & * & * \\ & 0 & * \\ & & 0 \end{pmatrix},$$

$$\mathfrak{b} = \begin{pmatrix} * & * & * \\ & * & * \\ & & * \end{pmatrix},$$

$$N = \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} * & * & * \\ & * & * \\ & & * \end{pmatrix}.$$

Using that the sum of two positive roots, if it is a root, is positive, we see that $\mathfrak{n}$ and $\mathfrak{b}$ are Lie subalgebras of $\mathfrak{g}$, with $\mathfrak{n}$ an ideal of $\mathfrak{b}$. Hence they correspond to some connected (complex) Lie subgroups N and B of G, with N a normal Lie subgroup of B.

Lemma 155. *(i) The Lie subgroups N, A and B of G are closed.*
(ii) $N \cap A = \{1\}$.
(iii) $B = NA$; more precisely, B is the semidirect product of N and A.

Proof. In the special case $K = \mathrm{U}(n)$, these conclusions are clear by inspection of the explicit descriptions given above (which depended upon the choice of C, but the choice doesn't matter, because any two choices are W-conjugate). In general, choose a faithful representation (π, V) of K and a basis $e_1, \ldots, e_n$ of V consisting of weight vectors for T, thus $te_j = t^{\lambda_j} e_j$ for some $\lambda_j \in X(T)$. We may assume that this basis has been ordered in such a way that if $\lambda_i < \lambda_j$ (with respect to the partial order defined by C), then $i < j$. Since for $x \in \mathfrak{g}^\alpha$ with $\alpha > 0$ we have $d\pi(x) : V^\lambda \to V^{\lambda+\alpha}$ and $\lambda + \alpha > \lambda$, we see that $d\pi(x) \in \begin{pmatrix} 0 & * & * \\ & 0 & * \\ & & 0 \end{pmatrix}$. On the other hand, $d\pi(\mathfrak{t}_{\mathbb{C}}) \subseteq \begin{pmatrix} * & * & * \\ & * & * \\ & & * \end{pmatrix}$. Exponentiating, it follows that $\pi(N) \subseteq \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix}$ and $\pi(A) \subseteq \begin{pmatrix} * & & \\ & * & \\ & & * \end{pmatrix}$. In particular, $A \cap N = \{1\}$.

We claim now that A and N are closed. (The remaining assertions then follow readily.) We identify G with its image under π. (Apologies in advance if these proofs are sloppy. The customary way to develop this material is in the language of algebraic groups. I don't know offhand of a reference for the approach taken here.)

Start with N. Recall that, by definition, N is generated by the image under exp of small elements of $\mathfrak{n}$. Let $\mathfrak{n}_1 := \begin{pmatrix} 0 & * & * \\ & 0 & * \\ & & 0 \end{pmatrix} \geq \mathfrak{n}$ denote the Lie algebra consisting of strictly upper-triangular matrices, and $N_1 := \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix} \geq N$ the corresponding Lie gruop. Then we have mutually inverse diffeomorphisms exp : $\mathfrak{n}_1 \to N_1$ and log : $N_1 \to \mathfrak{n}_1$ given by finite Taylor series. In particular, the exponential map is defined on all of $\mathfrak{n}$. Using the BCH formula, we see that in fact $N = \exp(\mathfrak{n})$. Since $\mathfrak{n}$ is closed in $\mathfrak{n}_1$, it follows that N is closed in N_1. Since $N_1 \cap G$ is closed in G, we conclude that N is closed in G.

We turn to A. Let $A_1 := \begin{pmatrix} \mathbb{C}^\times & & \\ & \mathbb{C}^\times & \\ & & \mathbb{C}^\times \end{pmatrix} \geq A$ denote the diagonal subgroup

of $\mathrm{GL}_n(\mathbb{C})$, and $T_1 := \begin{pmatrix} \mathrm{U}(1) & & \\ & \mathrm{U}(1) & \\ & & \mathrm{U}(1) \end{pmatrix}$ its standard maximal torus, so that

$A_1 \geq T_1 \geq T$. We get $\mathfrak{t}_{1,\mathbb{R}} \geq \mathfrak{t}_\mathbb{R}$ and $\mathfrak{t}_{1,\mathbb{Z}} \geq \mathfrak{t}_\mathbb{Z}$. Since $T_1 \hookrightarrow T$, we have $\mathfrak{t}_\mathbb{Z} = \mathfrak{t}_{1,\mathbb{Z}} \cap \mathfrak{t}_\mathbb{R}$. Thus we can find a lattice $L \subseteq \mathfrak{t}_{1,\mathbb{Z}}$ complementary to $\mathfrak{t}_\mathbb{Z}$, i.e., so that $\mathfrak{t}_{1,\mathbb{Z}} = \mathfrak{t}_\mathbb{Z} \oplus L$. Choosing bases for these lattices, we can identify $\mathfrak{t}_{1,\mathbb{Z}}$ with $\mathbb{Z}^n$ and $\mathfrak{t}_\mathbb{Z}$ with $\mathbb{Z}^m$ for some $0 \leq m \leq n$ in such a way that the inclusion $\mathfrak{t}_\mathbb{Z} \hookrightarrow \mathfrak{t}_{1,\mathbb{Z}}$ is the standard one. In the coordinates defined by these bases, the inclusions $T = \mathrm{U}(1)^m \hookrightarrow T_1 = \mathrm{U}(1)^n$ and $A = (\mathbb{C}^\times)^m \hookrightarrow (\mathbb{C}^\times)^n = A_1$ are then likewise the standard ones; in particular, the latter has closed image. $\qquad\square$

In particular, $B/N \cong A$. We may and shall thus extend each homomorphism $\chi : A \to \mathbb{C}^\times$ to a homomorphism $\chi : B \to \mathbb{C}^\times$ that is trivial on N, thus $\chi(na) = \chi(a)$. We'll often write $b^\chi := \chi(b)$.

Theorem 156. *Let π be an irreducible holomorphic finite-dimensional representation of G. There is a unique holomorphic character $\lambda : A \to \mathbb{C}^\times$ and a unique-up-to-scalars nonzero vector $v \in \pi$ so that*

$$\pi(b)v = b^\lambda v \text{ for all } b \in B. \qquad (90)$$

Moreover, λ is the highest weight of π and v is a highest weight vector.

Proof. Existence: Let $\lambda \in X(T)$ be the highest weight of π (via Theorem 153, and using Theorem 83 to regard π as an irreducible representation of K), and $v \in \pi[\lambda]$ a highest weight vector. We extend λ to a holomorphic character $A \to \mathbb{C}^\times$. The vector v is an eigenvector of T with eigenvalue λ, hence also of A with eigenvalue λ. A short calculation as in the proof of Lemma 92 shows that $d\pi(X_\alpha) : \pi[\lambda] \to \pi[\lambda+\alpha]$. Since λ is the highest weight of π, we have $\pi[\lambda+\alpha] = \{0\}$. Thus $X_\alpha v = 0$. Since α was arbitrary, v is annihilated by $\mathfrak{n}$. Since N is connected (by construction), it follows that v is fixed by N. Thus v is a B-eigenvector, as required.

Uniqueness: Suppose given a pair (v, λ) as in the conclusion of the theorem. Our task is to show that λ is in fact the highest weight π. For each finite ordered tuple $(\alpha_1, \ldots, \alpha_n)$ of positive roots, the vector

$$X_{-\alpha_1} \cdots X_{-\alpha_n} v := d\pi(X_{-\alpha_1}) \cdots d\pi(X_{-\alpha_n})v$$

is a T-eigenvector of weight $\lambda - \alpha_1 - \cdots - \alpha_n$, which is strictly lower than λ when $n \geq 1$. Call the span of the indicated vectors V. It suffices to show that $V = \pi$. Since $v \in V$ and π is irreducible, it suffices to show that V is K-invariant. Since K is connected, it is enough to show that V is $\mathfrak{g}$-invariant. To that end, recall that $\mathfrak{g} = \mathfrak{t}_\mathbb{C} \oplus (\oplus_{\alpha>0}\mathbb{C}X_\alpha) \oplus (\oplus_{\alpha>0}\mathbb{C}X_{-\alpha})$. It's clear that V is invariant by $\mathfrak{t}_\mathbb{C}$ and by $X_{-\alpha}$ for $\alpha > 0$, so it remains only to show for $\alpha > 0$ that V is X_α-invariant. Differentiating the condition (90), we see that v is annihilated by X_α. It remains to check for $n \geq 1$ that $X_\alpha X_{-\alpha_1} \cdots X_{-\alpha_n} v$ belongs to V. Set $u := X_{-\alpha_2} \cdots X_{-\alpha_n} v$. Then $X_\alpha X_{-\alpha_1} u = X_{-\alpha_1} X_\alpha u + [X_\alpha, X_{-\alpha_1}]u$. By induction on n, both $X_{-\alpha}u$ and $[X_\alpha, X_{-\alpha_1}]u$ belong to V. Since $X_{-\alpha_1}$ stabilizes V, the required conclusion follows. $\qquad\square$

Now let $\lambda \in X(T)$ be a dominant weight, and consider the vector space

$$F_\lambda := \{f : G \to \mathbb{C} : \text{holomorphic}, f(bg) = b^{-\lambda}f(g) \text{ for } b \in B, g \in G\}.$$

The group G (and hence also its subgroup K) acts on F_λ by right translation. We will show the following:

Theorem 157. *Let π_λ denote the irreducible representation of K of highest weight λ. Then $\pi_\lambda \cong F_\lambda^*$. Here F_λ^* denotes the dual representation of F_λ, regarded as a representation of K; alternatively, we may use Theorem 83 to extend π_λ to an algebraic representation of G, and the isomorphism is then of algebraic representations of G.*

We compute in passing the dual of π_λ. Note that if C is a Weyl chamber, then so is $-C$; since the Weyl group acts simply-transitively on the chambers, there is thus a unique $w_0 \in W$ for which $w_0(C) = -C$. (It is often called the *longest Weyl element*, with the notion of "longest" corresponding to C). For instance, if

$K = \mathrm{U}(n)$ and C is chosen in the usual way, then $w_0 = \begin{pmatrix} & & 1 \\ & \cdots & \\ 1 & & \end{pmatrix}$.

Lemma 158. $\pi_\lambda^* \cong \pi_{-w_0(\lambda)}$.

Proof. It's clear that π_λ^* is an irreducible, hence is isomorphic to π_μ for some dominant weight μ. We must verify that $\mu = -w_0(\lambda)$, or equivalently, that $\lambda = -w_0(\mu)$. We check first that $-\mu$ is the lowest weight of π_λ. From this it follows that $-w_0(\mu)$ is the highest weight of π_λ, and so $\lambda = -w_0(\mu)$, as required. $\square$

Thus Theorem 157 tells us that

$$\pi_\lambda \cong F_{-w_0(\lambda)}.$$

Proof of Theorem 157. Let F denote the space of all holomorphic functions $f : G \to \mathbb{C}$, and let $f \in F$. Using the identity principle for holomorphic functions on $G \cong K \times \mathfrak{p}$, we may identify f with its restriction to K, and then further with an element of the space $L^2(K)$, on which the group $K \times K$ acts by the regular representation $(k_1, k_2) \cdot f(x) = f(k_1^{-1}xk)$. We accordingly have an $L^2(K)$-decomposition $f = \sum_\mu f_\mu$, where $\mu \in X(T)$ runs over the dominant weights and f_μ belongs to the space $\mathcal{A}(\pi_\mu)$ of matrix coefficients of the representation π_μ. Let L and R denote respectively the left and right regular representations of K on $L^2(K)$. Since $\mathcal{A}(\pi_\mu) \cong \pi_\mu^* \otimes \pi_\mu$ is the π_μ-isotypic component (see §5.10) of $L^2(K)$ under R, we have $f_\mu = \dim(\pi_\mu)R(\overline{\chi_\mu})f$. Suppose now that $f \in F_\lambda$, i.e., that $L(b)f = b^\lambda f$ for all $b \in B$. Since R and L commute, it follows that $L(b)f_\mu = b^\lambda f_\mu$, i.e., that f_μ belongs to F_λ. If f_μ is nonzero, then we may find a nonzero vector $v \in \pi_\mu^*$ so that $bv = b^\lambda v$ for all $b \in B$. Then Theorem 156 implies that $\pi_\mu^* \cong \pi_\lambda$ and that $v \in \pi_\lambda$ is a highest weight vector. Thus

$$f \in (\pi_\lambda \otimes \pi_\lambda^*) \cap F_\lambda = \mathbb{C}v \otimes \pi_\lambda^*.$$

Thus $F_\lambda = \mathbb{C}v \otimes \pi_\lambda^*$, hence $F_\lambda \cong \pi_\lambda^*$. $\square$

9. Plancherel formula for complex reductive groups

With the remaining fewl ectures in the course, we aim to treat some of the simplest aspects of the representation theory of non-compact groups. We focus on

the case of complex reductive groups, which are simpler than real reductive groups for reasons to be explained later.

Let G be a unimodular group equipped with a Haar measure dg, and let $f : G \to \mathbb{C}$ be nice enough (e.g., smooth and compactly-supported in the case of a Lie group). We've seen in §5.9 that for G compact, one has the Fourier inversion formula

$$f(1) = \sum_{\pi \in \mathrm{Irr}(G)} \frac{\dim(\pi)}{\mathrm{vol}(G)} \, \mathrm{trace}(\pi(f)).$$

(We showed this when $\mathrm{vol}(G) = 1$, but the general case follows because $\pi(f)$ and $\mathrm{vol}(G)$ scale in the same way when one scales dg.) In the non-compact case, one seeks a formula of the shape

$$f(1) = \int_{\pi \in \hat{G}} \mathrm{trace}(\pi(f)) \, d\mu_P(\pi), \tag{91}$$

where now $\hat{G}$ denotes the set of isomorphism classes of irreducible unitary representations of G and $d\mu_P$ denotes a measure on $\hat{G}$, called *Plancherel measure*, which scales inversely to dg. For instance, the Fourier inversion formulas on the circle

$$f(0) = \sum_{n \in \mathbb{Z}} \hat{f}(n), \quad \hat{f}(n) := \int_{x \in \mathbb{R}/\mathbb{Z}} f(x) e^{2\pi i n x} \, dx, \quad f \in C^\infty(\mathbb{R}/\mathbb{Z})$$

and on the real line

$$f(0) = \sum_{\xi \in \mathbb{R}} \hat{f}(\xi) \, d\xi, \quad \hat{f}(\xi) := \int_{x \in \mathbb{R}/\mathbb{Z}} f(x) e^{2\pi i n x} \, dx, \quad f \in C_c^\infty(\mathbb{R})$$

may be interpreted in this way; for instance, in the second example, $\hat{f}(\xi)$ is the trace of $\pi_\xi(f)$, where π_ξ denotes the one-dimensional representation of $\mathbb{R}$, spanned by the function $\mathbb{R} \ni x \mapsto e^{2\pi i \xi x}$, and with $\mathbb{R}$ acting by right translation.

In a course emphasizing the functional analytic aspects of representation theory one might study general conditions under which such formulas (91) exist (e.g., look up the definition of a "Type I" group). Here we will focus instead on the problem of determining μ_P explicitly for a more restricted class of groups, namely the connected complex reductive groups.

We will use much of the same notation as above. We start with a compact connected Lie group K. We denote by G its complexification. We fix a maximal torus T and Weyl chamber C. We use these to define subgroups N, A and B of G, as in §8.3. The example $K = \mathrm{U}(n)$ depicted earlier is worth keeping in mind.

We note that G is unimodular. To see this, we must show that $|\det(\mathrm{Ad}(g))| = 1$ for all $g \in G$. It suffices to verify the stronger identity

$$\det(\mathrm{Ad}(g)) = 1 \tag{92}$$

By the identity principle for holomorphic functions applied to $\det \circ \mathrm{Ad} : G \to \mathbb{C}^\times$, we need only verify (92) holds for $g \in K$. Since $\det \circ \mathrm{Ad}$ is a class function, we reduce further to verifying (92) when $g = t \in T$, and indeed, $\det(\mathrm{Ad}(t)) = \prod_{\alpha \in \Phi} t^\alpha = \prod_{\alpha > 0} t^\alpha t^{-\alpha} = 1$.

9.1. **Iwasawa decomposition.** In its most basic form it asserts that

$$G = BK.$$

Moreover,

$$B \cap K = T.$$

In the case $G = \mathrm{GL}_n(\mathbb{C})$, this amounts to the Gram–Schmidt procedure (explanation given in class). We record the proof for general G. We consider first the analogous Lie algebra question. To start, we might write any element of $\mathfrak{g}$ as $x+y+z$ using the decomposition

$$\mathfrak{g} = \mathfrak{t}_{\mathbb{C}} \oplus \mathfrak{n} \oplus \bar{\mathfrak{n}},$$

where $\mathfrak{n} = \oplus_{\alpha>0}\mathfrak{g}^\alpha$ and $\bar{\mathfrak{n}} = \oplus_{\alpha<0}\mathfrak{g}^\alpha$. We may write $x = x_0 + x_1$ with $x_0 \in \mathfrak{t}_{\mathbb{R}}, x_1 \in i\mathfrak{t}_{\mathbb{R}}$, say. Then the element $w := x_1 + z + \theta(z)$ is θ-invariant, hence belongs to $\mathfrak{k}$. On the other hand, $y - \theta(z)$ belongs to $\mathfrak{n}$. Thus

$$x + y + z = w + x_0 + (y - \theta(z)) \in \mathfrak{k} + \mathfrak{t}_{\mathbb{R}} + \mathfrak{n}.$$

In particular,

$$\mathfrak{g} = \mathfrak{b} + \mathfrak{k}. \tag{93}$$

Using that $\mathfrak{k} = \{x \in \mathfrak{g} : \theta(x) = x\}$ and that θ swaps $\mathfrak{n}$ and $\bar{\mathfrak{n}}$ and acts on $\mathfrak{t}_{\mathbb{C}}$ with fixed subspace $I\mathfrak{t}_{\mathbb{R}}$, we see that

$$\mathfrak{b} \cap \mathfrak{k} = i\mathfrak{t}_{\mathbb{R}}.$$

We now turn to the groups. We've seen that B is a closed subgroup of G. Since K is compact, it follows that BK is closed. Using (93) and computing the derivative of the map $B \times K \xrightarrow{(b,g) \mapsto bg} G$ as in the proof of the Weyl integral formula, we see moreover that BK is open (see for instance Knapp, Lemma 5.11). Since G is connected, we conclude that $G = BK$.

9.2. **Integration on $B\backslash G$.** We fix a Haar measure dg on G and a left Haar measure db on B. Using these choices we will define a notion of integration over $B\backslash G$.

Let $\mathcal{F}$ denote the space of measurable functions $f : G \to \mathbb{C}$ such that

$$f(bg) = \delta(b)f(g)$$

for all $b \in B, g \in G$ and $\int_K |f| < \infty$, the latter integral taken with respect to any Haar measure on K.

Lemma 159. *Each $f \in \mathcal{F}$ is of the form $f(g) = \int_B \tilde{f}(bg)\, db$ for some $\tilde{f} : G \to \mathbb{C}$. The map $\mathcal{F} \ni f \mapsto \int_G \tilde{f}(g)\, dg$ is a well-defined linear functional on $\mathcal{F}$ that is invariant under right translation by G. We denote this linear functional by $f \mapsto \int_{B\backslash G} f$. We have $\int_{B\backslash G} f = c \int_K f$ for some constant $c > 0$, depending only upon the various choices of Haar measure.*

Proof. Left to the reader; this or some variant was treated in "Lie groups I". $\square$

9.3. **Principal series representations.** We denote by $A^\wedge$ the group of unitary characters of A, thus $A^\wedge$ consists of continuous homomorphisms $\chi : A \to \mathrm{U}(1)$. We similarly define the group $\mathfrak{a}^\wedge$ of unitary characters of A. Using the surjective quotient map $\mathfrak{a} \to A$ given by exponentiation, we may identify $A^\wedge$ with a subgroup of $\mathfrak{a}^\wedge$. On the other hand, by differentiating at the origin, we may identify $\mathfrak{a}^\wedge$ with the $\mathbb{R}$-vector space consisting of linear functionals $\mathfrak{a} \to i\mathbb{R}$. Concretely, suppose A is the group of diagonal matrices in $\mathrm{GL}_n(\mathbb{C})$. Write $a \in A$ in coordinates as $a = (a_1, \ldots, a_n)$ with $a_j \in \mathbb{C}^\times$. Then every $\chi \in A^\wedge$ is of the form

$\chi(a) = \prod_j a_j^{s_j}(a_j/|a_j|)^{k_j}$ for some $s_j \in i\mathbb{R}$ and $k_j \in \mathbb{Z}$, and this defines an isomorphism $A^{\wedge} \cong (i\mathbb{R})^n \times \mathbb{Z}^n$. Similarly, we may identify $\mathfrak{a}$ with the group $\mathbb{C}^n$ of diagonal matrices $x = (x_1, \ldots, x_n)$ in $M_n(\mathbb{C})$ and $\mathfrak{a}^{\wedge}$ with the group $(i\mathbb{R})^n \times \mathbb{R}^n$.

To each $\lambda \in \mathrm{Hom}(A, \mathbb{C}^{\times})$ we attach a "principal series representation" π_{λ}, consisting of measurable functions $v : G \to \mathbb{C}$ satisfying

$$v(bx) = \delta^{1/2}(b)b^{\lambda}v(x)$$

for all $b \in B, x \in G$ and

$$\int_K |v|^2 < \infty.$$

Thanks to the Iwasawa decomposition, we may regard π_{λ} as a closed subspace of the Hilbert space $L^2(K)$. The group G acts on π_{λ} by right translation. We note that if $f_1 \in \pi_{\lambda}$ and $f_2 \in \pi_{-\lambda}$, then $f_1 f_2$ belongs to the space $\mathcal{F}$ considered in §9.2, and so we obtain a G-invariant pairing

$$\pi_{\lambda} \otimes \pi_{-\lambda} \to \mathbb{C}$$

$$f_1 \otimes f_2 \mapsto \int_{B \backslash G} f_1 f_2.$$

Also, $\overline{\pi_{\lambda}} = \pi_{-\overline{\lambda}}$. Thus if λ belongs to $A^{\wedge} = \mathrm{Hom}(A, \mathrm{U}(1))$, so that $\lambda = -\overline{\lambda}$, then integration over $B \backslash G$ defines a G-invariant inner product on π_{λ}, hence π_{λ} is a unitary representation of G.

9.4. Characters of principal series representations.

Recall that we equip G with a Haar measure dg and B with a left Haar measure db; this normalizes an integral over $B \backslash G$ in the sense described above. To each $f \in C_c(G)$ we attach an integral operator $\pi(f) \in \mathrm{End}(\pi)$, as in §5, by the formula $\pi(f) := \int_{g \in G} f(g)\pi(g)\, dg$.

Theorem 160. *For each $f \in C_c^{\infty}(G)$, the operator $\pi_{\lambda}(f)$ is trace class. We have*

$$\mathrm{trace}(\pi_{\lambda}(f)) = \int_{g \in B \backslash G} \int_{b \in B} f(g^{-1}bg)\delta^{1/2}(b)b^{\lambda}. \tag{94}$$

Proof. Let $v \in \pi_{\lambda}$. By definition,

$$\pi(f)v(x) = \int_{g \in G} f(g)v(xg),$$

where here and henceforth we the omit Haar measures for notational simplicity. We apply the substitutinon $g \mapsto x^{-1}g$ to rewrite the above as

$$\int_{g \in G} f(x^{-1}g)v(g).$$

We next use the defining property of $\int_{B \backslash G}$ to rewrite the above as

$$\int_{y \in B \backslash G} \int_{b \in B} f(x^{-1}by)v(by).$$

Since $v(by) = \delta^{1/2}(b)b^{\lambda}v(y)$, it follows that

$$\pi(f)v(x) = \int_{y \in B \backslash G} v(y)k(x, y) \tag{95}$$

with

$$k(x, y) := \int_{b \in B} \delta^{1/2}(b)b^{\lambda}f(x^{-1}by).$$

Since B is closed and $f \in C_c^\infty(G)$, we have $k \in C^\infty(G \times G)$. It follows readily from the definition that

$$k(b_1 x, b_2 y) = \delta^{1/2}(b_1) b_1^\lambda \delta^{1/2}(b_2) b_2^{-\lambda} k(x, y) \text{ for } b_1, b_2 \in B. \tag{96}$$

Note that the diagonal kernel integral

$$\int_{x \in B \backslash G} k(x, x) \tag{97}$$

is the RHS of (94).

The remainder of the argument is a bit of standard functional analysis. It suffice to show more generally that if T is any operator on π defined as in (95) by a smooth kernel $k \in C^\infty(G \times G)$ satisfying (96), then T defines a trace class operator with trace given by (97). This can be proved in the same way as the following statement, which may have come up in a functional analysis course:

Let (M, μ) be a compact manifold equipped with a smooth finite measure, and let $k \in C^\infty(M \times M)$. Then the operator T on $L^2(M, \mu)$ defined by $Tv(x) := \int v(y) k(x, y) \, d\mu(y)$ is trace class, with $\text{trace}(T) = \int k(x, x) \, d\mu(x)$.

One can prove this by reducing first to the case that M is a compact torus and then appealing to Fourier series.

$\square$

9.5. **Some basics concerning conjugacy classes in G.** Every such class intersects B, that is to say, $G = \cup_{x \in B \backslash G} x^{-1} B x$. For instance, if $G = \text{GL}_n(\mathbb{C})$, then this follows from the Jordan normal form. To sketch a proof in general, it suffices to show for each $g \in G$ that the map

$$g : B \backslash G \to B \backslash G$$

$$Bx \mapsto Bxg$$

has a fixed point (because if $Bx = Bxg$, then $g \in x^{-1} B x$). One way to see this is to use that $B \backslash G \cong T \backslash K$ (a consequence of the Iwasawa decomposition) and then to show as in one of the standard proofs of the conjugacy of maximal tori (not the one presented in this course) that any map $T \backslash K \to T \backslash K$ homotopic to the identity has a fixed point. One can also argue algebraically using the Borel fixed point theorem.

9.6. **A change of variables.** The groups N and A are unimodular (indeed, their adjoint representations are unimodular, hence have trivial determinant). We may thus equip them with Haar measures, compatibly with the chosen left Haar measure on B in the sense that

$$\int_B h = \int_{n \in N, a \in A} h(an) = \int_{n \in N, a \in A} h(na) \delta^{-1}(a).$$

For instance, if $G = \text{SL}_2(\mathbb{C})$, then a left Haar on B is given by

$$\int_B h = \int_{x \in \mathbb{C}, y \in \mathbb{C}^\times} h(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix}) \frac{dx \, d^\times y}{|y|_\mathbb{C}^2},$$

where dx and $d^\times y$ denote Haar measures on $\mathbb{C}$ and $\mathbb{C}^\times$ respectively and where

$$|y|_\mathbb{C} := |y|^2$$

denotes the complex modulus, given by the square of the usual complex absolute value, so that $\int_{x \in \mathbb{C}} h(x) \, dx = |y|_{\mathbb{C}} \int_{x \in \mathbb{C}} h(yx) \, dx$ for any $h \in C_c(\mathbb{C})$, and so that $d^\times y = dy/|y|_{\mathbb{C}}$ defines a Haar measure on $\mathbb{C}^\times$ for each Haar measure dy on $\mathbb{C}$.

We want to rewrite this integral formula for B so that it interacts better with conjugacy classes. To start, note that

$$\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} = \begin{pmatrix} y & xy^{-1} \\ & y^{-1} \end{pmatrix}.$$

Now observe that "generically" (i.e., for $y \neq 1$), the above matrix has distinct eigenvalues, hence is conjugate to the diagonal matrix with the same entries. More explicitly, we have

$$\begin{pmatrix} 1 & z \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} \begin{pmatrix} 1 & z \\ & 1 \end{pmatrix}^{-1} = \begin{pmatrix} y & (y^{-1} - y)z \\ & y^{-1} \end{pmatrix},$$

which equals $\begin{pmatrix} y & xy^{-1} \\ & y^{-1} \end{pmatrix}$ when $z = xy^{-1}/(y^{-1} - y)$. Thus

$$\int_{x \in \mathbb{C}} h\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} \right) = \frac{|y^{-1} - y|_{\mathbb{C}}}{|y^{-1}|_{\mathbb{C}}} \int_{z \in \mathbb{C}} h\left(\begin{pmatrix} 1 & z \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} \begin{pmatrix} 1 & z \\ & 1 \end{pmatrix}^{-1} \right) \tag{98}$$

and so

$$\int_B h = \frac{|y^{-1} - y|_{\mathbb{C}}}{|y|_{\mathbb{C}}} \int_{x \in \mathbb{C}} h\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}^{-1} \right) dx \, d^\times y.$$

Here it is understood that we integrate away from the measure zero set of y for which $y^2 = 1$.

A similar argument applied "one root at a time" gives the following more general identity:

Lemma 161. *With Haar measures on B, N, A compatibly normalized as above, we have for $h \in C_c(B)$*

$$\int_B h = \int_{a \in A^{\mathrm{reg}}} \frac{\prod_{\alpha > 0} |a^{\alpha/2} - a^{-\alpha/2}|_{\mathbb{C}}}{\delta^{1/2}(a)} \int_{n \in N} h(nan^{-1}),$$

with the integral taken over the full measure subset

$$A^{\mathrm{reg}} := \{a \in A : a^\alpha \neq 1 \text{ for all } \alpha \in \Phi\}.$$

Proof. Compare with Knapp, Lemma 10.16. The key point is to verify the following general form of (98), valid for any $a \in A^{\mathrm{reg}}$:

$$\int_{n \in N} h(na) = \frac{\prod_{\alpha > 0} |a^{\alpha/2} - a^{-\alpha/2}|_{\mathbb{C}}}{\delta^{-1/2}(a)} \int_{n \in N} h(nan^{-1}). \tag{99}$$

$\square$

As a consequence, we can rewrite the formula given above for χ_λ, as follows. We first define $F : A^{\mathrm{reg}} \to \mathbb{C}$ by the formula

$$F(a) := \left(\prod_{\alpha > 0} |a^{\alpha/2} - a^{-\alpha/2}|_{\mathbb{C}} \right) \int_{g \in A \backslash G} f(g^{-1}ag). \tag{100}$$

By working backwards through the above calculations, we see that F extends to an element of $C_c^\infty(A)$; indeed, by (99), we have for $a \in A^{\mathrm{reg}}$

$$F(a) = (\prod_{\alpha > 0} |a^{\alpha/2} - a^{-\alpha/2}|_{\mathbb{C}}) \int_{g \in B \backslash G} \int_{n \in N} f(g^{-1}nan^{-1}g)$$

$$= \delta^{-1/2}(a) \int_{g \in B \backslash G} \int_{n \in N} f(g^{-1}nag)$$

$$= \delta^{-1/2}(a) \int_{k \in K} \int_{n \in N} f(k^{-1}nag)$$

for a suitable Haar measure on K, but this last expression defines for $a \in A$ an element of $C_c^\infty(A)$, using here that K is compact. We define the Fourier transform $F^\wedge : A^\wedge \to \mathbb{C}$ by

$$F^\wedge(\lambda) := \int_{a \in A} a^\lambda F(a).$$

We then obtain

$$\chi_\lambda(f) = F^\wedge(\lambda).$$

We equip $A^\wedge$ with the Haar measure dual to that on A, so that the Fourier inversion formula $F(1) = \int_{\lambda \in A^\wedge} F^\wedge(\lambda)$ holds.

9.7. **Passage to the Lie algebra.** Suppose now that $f \in C_c^\infty(G)$ as above is supported in a "small conjugation-invariant neighborhood" U of the identity element 1 of the Lie group G. For example, in the case $G = \mathrm{GL}_n(\mathbb{C})$, we might define U to be the set of all $g \in G$ each of whose eigenvalues c satisfies $|c - 1| < \varepsilon$ for some small $\varepsilon > 0$; alternatively, we might ask that the subleading coefficients of the characteristic polynomial of g be bounded by ε. We may then find a small $\mathrm{Ad}(G)$-invariant neighborhood Y of the origin 0 of the Lie algebra $\mathfrak{g}$ such that $\exp : Y \to U$ is a diffeomorphism onto its image. We may define an element $f_0 \in C_c^\infty(Y)$ by requiring that $f_0(x) = f(\exp(x))$. We define $F_0 \in C_c^\infty(\mathfrak{a})$ to vanish off elements $x \in \mathfrak{a} \cap Y$, for which we set $F_0(x) := F(\exp(x))$, with F attached to f as above. We equip $\mathfrak{a}$ with the Haar measure matching up near the origin under the exponential map with the Haar measure on A near the identity. For $\lambda \in \mathfrak{a}^\wedge$ (hence, in particular, for $\lambda \in A^\wedge$), we may define the Fourier transform $F_0^\wedge(\lambda)$; we have $F_0^\wedge(\lambda) = F^\wedge(\lambda)$ when $\lambda \in A^\wedge$. (Recall from before that $\mathfrak{a}^\wedge \cong (\mathbb{R} \times \mathbb{R})^n$ while $A^\wedge \cong (\mathbb{R} \times \mathbb{Z})^n$.) We equip $\mathfrak{a}^\wedge$ with the dual Haar measure, so that the Fourier inversion formula $F_0(0) = \int_{\lambda \in \mathfrak{a}^\wedge} F_0^\wedge(\lambda)$ holds.

Recall that we've fixed a Haar measure dg on G. We equip $\mathfrak{g}$ with the Haar measure dx so that for small $x \in \mathfrak{g}$, writing $g = \exp(x)$, we have $dg = j(x)\, dx$, with $j(0) = 1$. Using Lemma 80, one may check then that the inverse square root of j admits the following formula for small enough regular elements $x \in \mathfrak{a}$:

$$j^{-1/2}(x) = \prod_{\alpha > 0} \left| \frac{e^{\alpha(a)/2} - e^{-\alpha(a)/2}}{\alpha(x)} \right|.$$

9.8. **Main result.** We define a polynomial function $P : \mathfrak{a}^\wedge \to \mathbb{R}$ by the formula $P(\lambda) := \prod_{\alpha > 0} |\lambda(H_\alpha)|_{\mathbb{C}}$. We aim to show the following:

Theorem 162. *For $f \in C_c^\infty(G)$, and with measures as normalized above,*

$$f(1) = \int_{\lambda \in A^\wedge} \frac{P(\lambda)}{|W|} \chi_\lambda(f).$$

This is a theorem of Gelfand–Naimark. We'll loosely follow Harish-Chandra's proof technique, as exposed in the textbook of Varadarajan. It is a fact that we will not have time to prove that the π_λ are *irreducible* unitary representations of G, so this theorem gives the desired Plancherel formula.

Let's suppose first that f is supported in a small conjugation-invariant neighborhood of 1, in the sense described above, so that we may define F, f_0, F_0. The formula (100) implies that for small regular $x \in \mathfrak{a}$, we have

$$F_0(x) = (\prod_{\alpha > 0} e^{\alpha(x)/2} - e^{-\alpha(x)/2}) \int_{g \in G/B} \int_{n \in N} f_0(\mathrm{Ad}(gn)x).$$

The exponential map restricts to a polynomial diffeomorphism $\exp : \mathfrak{n} \to N$ with polynomial inverse $\log : N \to \mathfrak{n}$. We may normalize the Haar measure on $\mathfrak{n}$ to be compatible with $\exp$. With the abbreviation $f_0^g(y) := f_0(\mathrm{Ad}(g)y)$, we then obtain

$$F_0(x) = (\prod_{\alpha > 0} e^{\alpha(x)/2} - e^{-\alpha(x)/2}) \int_{g \in G/B} \int_{y \in \mathfrak{n}} f_0^g(\mathrm{Ad}(e^y)x).$$

By the Lie algebra analogue of Lemma 161, we have

$$\int_{y \in \mathfrak{n}} f_0^g(\mathrm{Ad}(e^y)x) = (\prod_{\alpha > 0} |\alpha(x)|_{\mathbb{C}}^{-1}) \int_{y \in \mathfrak{n}} f_0^g(y + x). \tag{101}$$

Thus

$$F_0(x) = j^{-1/2}(x) \int_{g \in G/B, y \in \mathfrak{n}} f_0^g(x + y).$$

The function $j^{-1/2}$ extends to an $\mathrm{Ad}(G)$-invariant function on $\mathfrak{g}$. Setting $\phi := j^{-1/2} f_0$, we may rewrite the above formula as

$$F_0(x) = \int_{g \in G/B, y \in \mathfrak{n}} \phi^g(x + y).$$

Recall the decomposition $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \bar{\mathfrak{n}}$. Fix $K \hookrightarrow \mathrm{U}(n)$ and hence $G \hookrightarrow \mathrm{GL}_n(\mathbb{C})$ as usual. We then get a perfect $\mathrm{Ad}(G)$-equivariant pairing $\mathfrak{g} \otimes \mathfrak{g} \to \mathbb{C}$ given by $(x, y) \mapsto \mathrm{trace}(xy)$. Using this we may identify $\mathfrak{g}^\wedge$ with $\mathfrak{g}$. Under this identification, $\mathfrak{n}$ and $\bar{\mathfrak{n}}$ are dual to one another, while $\mathfrak{a}$ is dual to itself. Integrating the above formula for F_0 over regular $x \in \mathfrak{a}$ thus gives for all $\lambda \in \mathfrak{a}^\wedge$ that

$$F_0^\wedge(\lambda) = \int_{g \in G/B} \int_{\eta \in \mathfrak{b}^\perp \subseteq \mathfrak{g}^\wedge} \hat{\phi}^g(\lambda + \eta).$$

Here $\hat{\phi}$ denotes the Fourier transform and $\mathfrak{b}^\perp$ the subspace of $\mathfrak{g}^\wedge$ that annihilates $\mathfrak{b}$; by the above discussion, it identifies with $\mathfrak{n}$.

We say that $\lambda \in \mathfrak{a}^\wedge$ is *regular* if $\lambda(H_\alpha) \neq 0$ for all roots α. By the analogue of (101) for $\mathfrak{g}^\wedge$, we see that for regular $\lambda \in \mathfrak{a}^\wedge$,

$$F_0^\wedge(\lambda) = (\prod_{\alpha > 0} |\lambda(H_\alpha)|_{\mathbb{C}}) \int_{g \in G/A} \hat{\phi}(\mathrm{Ad}(g)\lambda).$$

Let's denote the latter expression by $\int_{\mathcal{O}_\lambda} \hat{\phi}$; it's a normalized integral over the G-orbit $\mathcal{O}_\lambda$ of λ. Then we have the following analogue of the Weyl integral formula for $h \in C_c(\mathfrak{g}^\wedge)$:

$$\int_{\mathfrak{g}^\wedge} h = \frac{1}{|W|} \int_{\lambda \in \mathfrak{a}^\wedge, \mathrm{regular}} P(\lambda) \int_{\mathcal{O}_\lambda} h.$$

(The key point here is that the union of the orbits $\mathcal{O}_\lambda$ is open and dense in G; we may thus reduce to the usual Jacobian calculation.)

We're now ready to prove the theorem. We denote by ∂_P the translation-invariant differential operator on $C_c^\infty(A)$ or on $C_c^\infty(\mathfrak{a})$ with symbol P, so that $(\partial_P h)^\wedge = P \cdot \hat{h}$. Let $f \in C_c^\infty(G)$. Define F as above. Then

$$\int_{\lambda \in \mathfrak{a}^\wedge} \frac{P(\lambda)}{|W|} \underbrace{\chi_\lambda(f)}_{=F^\wedge(\lambda)} = \frac{(\partial_P F)(1)}{|W|}.$$

We want to show that this last expression equals $f(1)$. By smoothly and $\mathrm{Ad}(G)$-invariantly truncating, we may assume that f is supported in a small conjugation-invariant neighborhood of the identity, so that we may define f_0, F_0, ϕ as above. Then, combining everything we've shown thus far, we conclude that

$$\begin{aligned}
\frac{(\partial_P F)(1)}{|W|} &= \frac{\partial_P F_0(0)}{|W|} \\
&= \int_{\lambda \in \mathfrak{a}^\wedge} \frac{P(\lambda)}{|W|} \int_{\mathcal{O}_\lambda} \hat{\phi} \\
&= \int_{\mathfrak{g}^\wedge} \hat{\phi} \\
&= \phi(0) \\
&= f_0(0) \\
&= f(1),
\end{aligned}$$

as required.

REFERENCES