The size of infinite-dimensional representations

David A. Vogan, Jr.
2-355, Department of Mathematics
MIT, Cambridge, MA 02139

This paper is offered in honor and in fond remembrance of Professor Bertram Kostant

March 16, 2017

Abstract

An infinite-dimensional representation π of a real reductive Lie group G can often be thought of as a function space on some manifold X. Although X is not uniquely defined by π, there are "geometric invariants" of π, first introduced by Roger Howe in the 1970s, related to the geometry of X. These invariants are easy to define but difficult to compute. I will describe some of the invariants, and recent progress toward computing them.

1 Introduction

One of the most basic questions we ask about any mathematical object is *how big is it?* In the case of finite sets, we ask how many elements. In the case of an enormous variety of infinite sets, we ask about *dimension*: for vector spaces, manifolds, rings, algebraic varieties...

When we use a group G to study almost any kind of mathematics, a fundamental tool is the collection of *representations* of G: homomorphisms

$$\pi\colon G \to GL(V) \tag{1.1}$$

from G to the group of invertible linear transformations of a vector space V. A nice example to keep in mind is

$$\begin{aligned} K &= SO(3) = \text{rotations of } \mathbb{R}^3 \\ V(\lambda) &= \{f \in C^\infty(S^2) \mid \Delta f = \lambda f\}, \end{aligned} \tag{1.2}$$

the λ-eigenspace of the Laplacian on the two-sphere. Because the Laplacian is elliptic and the sphere is compact, this is a finite-dimensional space; in fact

$$\dim V(\lambda) = \begin{cases} 2k+1, & \lambda = -k(k+1) \quad (k \in \mathbb{N}) \\ 0 & \lambda \neq -k(k+1). \end{cases}$$

It's also true that $V(-k(k+1))$ is an irreducible representation of K, and that every irreducible representation is of this form. This is a beautiful result, with lots of applications in mathematics and physics. The applications arise in part because any nice function on the sphere can be written as an (infinite) linear combination of these eigenfunctions.

For a slightly different example, let

$$\begin{aligned} A &= \mathbb{R}^2 = \text{translations of } \mathbb{R}^2 \\ E(\mu_1, \mu_2) &= \{ f \in C^\infty(\mathbb{R}^2) \mid \frac{\partial f}{\partial x_i} = \mu_i f \}. \end{aligned} \tag{1.3}$$

Then

$$E(\mu) = \mathbb{C} \cdot e^{\mu_1 x_1 + \mu_2 x_2}, \qquad \dim E(\mu) = 1.$$

Again these eigenspaces are all irreducible representations of A; their importance arises in part from the possibility of expressing arbitrary nice functions on the plane as (continuous) linear combinations of these eigenfunctions.

An important special feature of the first example is that K is compact; and of the second, that A is abelian. If G is neither compact nor abelian, then interesting representations of G are typically infinite-dimensional. Here is an example. Consider

$$\begin{aligned} G &= SO(2,1) \\ &= \text{determinant one linear maps preserving } Q(x) = x_1^2 + x_2^2 - x_3^2, \\ X &= \text{hyperboloid of two sheets } = \{ x \in \mathbb{R}^3 \mid Q(x) = -1 \} \\ I(\tau) &= \{ f \in C^\infty(X) \mid \Delta f = \tau f \}; \end{aligned} \tag{1.4}$$

here Δ is the Laplace operator for the G-invariant Riemannian metric on X. The map $x \mapsto -x$ exchanges the two connected components of X and commutes with the action of G and Δ; so

$$I(\tau) = I(\tau)_{\text{even}} \oplus I(\tau)_{\text{odd}}; \tag{1.5}$$

these two spaces are identified by restriction to a single component of X.

The Laplace operator Δ is still elliptic, which guarantees that the functions in $I(\tau)$ are real analytic; but X is noncompact, so there is no reason for $I(\tau)$ to be finite-dimensional. Indeed $I(\tau)$ is infinite-dimensional for every τ. For $\tau = 0$, $I(0)_{\text{even}}$ has an obvious one-dimensional G-invariant subspace consisting of the constant functions. Not quite so obviously, the three-dimensional space of (restrictions to X of) linear functions is a G-invariant subspace of $I(2)_{\text{odd}}$. For every nonnegative integer k, there is a $(2k+1)$-dimensional G-invariant subspace

$$J_{2k+1} \subset I(k^2 + k)_{(k \mod 2)} \tag{1.6}$$

consisting of the restrictions to X of certain polynomials of degree k. These are essentially the *only* closed G-invariant subspaces of $I(\tau)$:

Proposition 1.7. *Suppose $G = SO(2,1)$ is the indefinite orthogonal group in three variables, X is the hyperboloid of two sheets, $\tau \in \mathbb{C}$ is complex, and*

$$\epsilon = \textit{even or odd}$$

is a parity.

1. *If $\tau \neq k^2+k$ for k a nonnegative integer, then the eigenspace $I(\tau)_\epsilon$ defined in* (1.4) *of is an irreducible representation of $SO(2,1)$.*
2. *If k is a nonnegative integer, then $I(k^2+k)_\epsilon$ has a finite-dimensional G-invariant subspace $J_{2k+1,\epsilon}$ of dimension $2k+1$. The quotient representation $I(k^2+k)_\epsilon/J_{2k+1,\epsilon}$ is irreducible.*
3. *Every irreducible quasisimple (see* (4.2i) *below) representation of G is infinitesimally equivalent either to $I(\tau)_\epsilon$ (for $\tau \neq k^2+k$), or to $J_{2k+1,\epsilon}$, or to $I(k^2+k)_\epsilon/J_{2k+1,\epsilon}$.*
4. *All of these representations are inequivalent, except that*

$$I(k^2+k)_{even}/J_{2k+1,even} \simeq I(k^2+k)_{odd}/J_{2k+1,odd}$$

The moral of these examples is that one can hope to construct representations of a Lie group G on (simultaneous) eigenspaces of (families of) G-invariant differential operators.

The question we wish to address is this: *how large is a particular infinite-dimensional representation of G?*

The example of $SO(2,1)$ shows that there are at least two aspects to this question. First, how large is an eigenspace representation of G? That is, how many solutions are there to interesting G-invariant simultaneous eigenvalue problems? Second, how can we keep track of the sizes of pieces when these eigenspace representations are reducible?

We will consider here only representations of a real reductive Lie group G, such as $SO(2,1)$ or $GL(n,\mathbb{R})$. We will introduce these groups in Section 2, and recall the representation theory of compact Lie groups in Section 3. In Section 4 we recall Harish-Chandra's basic results about the infinite-dimensional representations of real reductive Lie groups. This leads to the definition of *Gelfand-Kirillov dimension* in Proposition 4.6, which is the simplest notion of size for infinite-dimensional representations. In Theorem 6.6 we will realize the Gelfand-Kirillov dimension as the dimension of an algebraic variety attached to the representation. Finally, in Section 9 we will talk about how one can compute this algebraic variety.

2 Real reductive groups

In this section I'll describe a class of Lie groups for which there is a very nice way to talk about the size of infinite-dimensional representations.

We begin with the general linear group

$$\begin{aligned} G_n &= GL(n,\mathbb{R}) = \text{invertible } n\times n \text{ real matrices},\\ \mathfrak{g}_n &= \mathrm{Lie}(G) = \mathfrak{gl}(n,\mathbb{R}) = n\times n \text{ real matrices}. \end{aligned} \tag{2.1a}$$

A fundamental structural fact is the *polar decomposition*: every $g \in G_n$ can be written uniquely as

$$g = kp, \qquad k \text{ orthogonal, } p \text{ symmetric positive definite.} \tag{2.1b}$$

If we write

$$\begin{aligned} K_n &= O(n) = \text{orthogonal group},\\ S_n &= \text{positive definite symmetric matrices} \end{aligned} \tag{2.1c}$$

then we have a smooth product decomposition (as manifolds)

$$G_n = K_n S_n. \tag{2.1d}$$

Writing

$$\mathfrak{s}_n = \text{real } n\times n \text{ symmetric matrices}, \tag{2.1e}$$

we see that

$$\exp\colon \mathfrak{s}_n \to S_n \tag{2.1f}$$

is a diffeomorphism, so we get a diffeomorphism

$$G_n \simeq K_n \times \mathfrak{s}_n, \qquad (k,X) \mapsto k\exp(X) \qquad (k\in K_n, X\in\mathfrak{s}_n). \tag{2.1g}$$

All of this structure is connected to the *Cartan involution*

$$\theta_n(g) = {}^t g^{-1} \quad (g\in G_n), \qquad \theta_n(X) = -{}^t X \quad (X \in \mathfrak{g}_n). \tag{2.1h}$$

For example,

$$\begin{aligned} K_n &= (G_n)^{\theta_n} = \{g\in G_n \mid \theta_n(g) = g\},\\ S_n &\subset \{g \in G_n \mid \theta_n(g) = g^{-1}\}. \end{aligned} \tag{2.1i}$$

It follows easily that the polar decomposition $g = ks$ may be computed by

$$s = \text{(unique) positive square root of } \theta(g)^{-1}g, \qquad k = gs^{-1}. \tag{2.1j}$$

The polar decomposition exhibits the manifold $G_n = GL(n,\mathbb{R})$ as the product of the compact group $K_n = O(n)$ and the vector space $\mathfrak{s}_n$. A familiar consequence is that many topological questions about G_n reduce immediately and completely to K_n. What matters for the representation theory of G_n is that *many interesting differential equations on G_n have unique solutions specified by initial conditions on K_n.*

The polar decomposition is inherited by many subgroups of $GL(n,\mathbb{R})$.

Proposition 2.2 (Cartan decomposition)**.** *Suppose G is a closed subgroup of $GL(n,\mathbb{R})$. Assume that*

1. $\theta_n(G) = G$ (that is, that G is closed under transpose of matrices), and

2. G has finitely many connected components.

Define

$$\theta = \theta_n|_G, \quad K = G^\theta = G \cap O(n),$$
$$S = G \cap S, \quad \mathfrak{s} = \mathfrak{g}^{-\theta} = \mathfrak{g} \cap \mathfrak{s}_n.$$

Then the polar decomposition for $GL(n,\mathbb{R})$ respects the subgroup G; so

$$G \simeq K \times \mathfrak{s}, \qquad (k, X) \mapsto k\exp(X) \qquad (k \in K, X \in \mathfrak{s}).$$

The "natural" definition of a real reductive Lie group is somewhat technical and complicated. But the polar decomposition is such a fundamental property that we can use it as the definition.

> **Definition 2.3.** (Knapp [7, page 3]) A *real reductive Lie group* is a closed subgroup $G \subset GL(n,\mathbb{R})$ having finitely many connected components and preserved by matrix transpose. Write
>
> $$\theta = \theta_n|_G, \qquad K = G^\theta = G \cap O(n),$$
> $$\mathfrak{g} = \mathrm{Lie}(G) \subset \mathfrak{gl}(n,\mathbb{R})$$
> $$S = G \cap S, \quad \mathfrak{s} = \mathfrak{g}^{-\theta} = \mathfrak{g} \cap \mathfrak{s}_n.$$
>
> The involution θ is called the *Cartan involution of G*. As a consequence of the *Cartan decomposition*
>
> $$G \simeq K \times \mathfrak{s}, \qquad (k, X) \mapsto k\exp(X)$$
>
> of Proposition 2.2, the group K is a maximal compact subgroup of G.

If G is the group of real points of a complex reductive algebraic group, then it can be shown that G is a real reductive group in the sense of this definition. This fact is what makes the terminology legal; but what makes it useful is that (for example) all of the classical reductive matrix groups like $SO(p,q)$ and $Sp(2n,\mathbb{R})$ are (in some standard realization) subgroups of $GL(N,\mathbb{R})$ preserved by transpose.

3 Representations of compact Lie groups

Our goal is to say something about the size of infinite-dimensional representations of real reductive Lie groups. The tool that we'll use is the representations of compact Lie groups; so in this section we'll recall some background facts about that.

Suppose therefore that K is a compact Lie group. A *representation* of K is a complete locally convex topological vector space E over $\mathbb{C}$, equipped with a homomorphism

$$\tau \colon K \to \mathrm{Aut}(E) \tag{3.1a}$$

with the property that the map

$$K \times E \to E, \qquad (k, e) \mapsto \tau(k)e$$

is continuous. A *subrepresentation* is a closed subspace $F \subset E$ with the property that

$$\tau(K)F \subset F. \tag{3.1b}$$

We say that (τ, E) is *irreducible* if there are exactly two subrepresentations (which must then be equal to 0 and E, with $E \neq 0$). There is an obvious notion of equivalence of representations, and we define

$$\widehat{K} = \text{equivalence classes of irreducible representations of } K. \tag{3.1c}$$

It turns out that $\widehat{K}$ is a countable set, consisting of finite-dimensional irreducible representations. We want to say a little about the structure of this set.

A *maximal torus* in K is a maximal closed connected abelian subgroup

$$T \subset K, \qquad \mathfrak{t} = \mathrm{Lie}(T) \subset \mathrm{Lie}(K) = \mathfrak{k}. \tag{3.1d}$$

Every irreducible representation of T has dimension 1 and takes values in $U(1) \subset GL(1, \mathbb{C})$; so

$$\widehat{T} = \mathrm{Hom}(T, U(1)) \simeq \{\lambda \in i\mathfrak{t}^* \mid \lambda(\ker \exp) \subset 2\pi i \mathbb{Z}\}; \tag{3.1e}$$

the isomorphism arises by taking the differential of a character. The last set is easily seen to be a lattice in the real vector space $i\mathfrak{t}^*$; we call it $X^*(T)$, the *character lattice of* T.

For $\lambda \in X^*(T)$, we write

$$\mathbb{C}_\lambda = \text{one-dimensional representation of } T. \tag{3.1f}$$

If E is any finite-dimensional representation of T, then we can define

$$m_E \colon X^*(T) \to \mathbb{N}, \qquad m_E(\lambda) = \dim \mathrm{Hom}_T(\mathbb{C}_\lambda, E), \tag{3.1g}$$

the *weight multiplicity function of* E. There is an isomorphism of representations

$$E \simeq \sum_{\lambda \in X^*(T)} \mathbb{C}_\lambda^{m_E(\lambda)}. \tag{3.1h}$$

Therefore *the multiplicity function determines the representation of* T.

So the representation theory of the connected abelian group T is quite simple: an irreducible representation is a point in a lattice, and a possibly reducible representation is an $\mathbb{N}$-valued function on $X^*(T)$ with finite support.

We want to use this information about T to (approximately) describe the (much more complicated) irreducible representations of K. One reason for speaking in some detail about this (very classical) process is that we are going to imitate it later: we will (approximately) describe the irreducible representations of a real reductive group G in terms of the representation theory of the maximal compact subgroup K.

Suppose now that (τ, E) is an irreducible (and therefore finite-dimensional) representation of K:

$$(\tau, E) \in \widehat{K}. \tag{3.1i}$$

We can attach to τ its *weight multiplicity function*

$$m_E = m_\tau \colon X^*(T) \to \mathbb{N}. \tag{3.1j}$$

The *weights* of τ are the characters $\lambda \in X^*(T)$ for which $m_\tau(\lambda) > 0$.

We are going to describe $\widehat{K}$ using weight multiplicity functions. In order to do that, we need one more definition. Set

$$\begin{aligned} T_1 &= Z_K(T) = \text{centralizer in } K \text{ of } T\\ N &= N_K(T) = \text{normalizer in } K \text{ of } T\\ W &= N/T_1 = \text{Weyl group of } T \text{ in } K \subset \mathrm{Aut}(T). \end{aligned} \tag{3.1k}$$

Because T is a *maximal* torus, T_1/T is discrete; since T_1 is obviously compact, T_1/T is finite. The automorphism group

$$\mathrm{Aut}(T) = \mathrm{Aut}(X^*(T)) = \mathrm{Aut}(\widehat{T}) \tag{3.1l}$$

is discrete; since N is obviously compact, W must be finite as well. We may therefore choose a negative definite inner product

$$\langle , \rangle \colon \mathfrak{t} \times \mathfrak{t} \to \mathbb{R} \tag{3.1m}$$

that is preserved by W. One way to do this is to choose an embedding $K \subset O(n)$ (that is, to identify K as a real reductive Lie group) and to define

$$\langle X, Y \rangle = \mathrm{tr}(XY) \qquad (X, Y \in \mathfrak{t}). \tag{3.1n}$$

The reason this inner product is negative definite is that the square of a nonzero skew-symmetric matrix must have strictly negative trace.

Our chosen inner product defines by duality a negative-definite inner product on $\mathfrak{t}^*$, and therefore a positive-definite inner product on

$$\langle , \rangle \colon i\mathfrak{t}^* \supset X^*(T) \to \mathbb{R}, \qquad \langle w \cdot \lambda, w \cdot \mu \rangle = \langle \lambda, \mu \rangle. \tag{3.1o}$$

Once we have made this choice, we define an *extremal weight* of τ to be a weight of maximal length in our chosen inner product.

Theorem 3.2 (Cartan-Weyl)**.** *Suppose $T \subset K$ is a maximal torus in a compact Lie group, with Weyl group $W = W(K, T) \subset \mathrm{Aut}(T)$ as in (3.1). If τ is an irreducible representation of K, then the set of extremal weights of τ is a single orbit of W, and is independent of the choice of W-invariant inner product. The extremal weight map*

$$\widehat{K} \to X^*(T)/W$$

defined in this way is surjective, with finite fibers; if K is connected, the map is bijective.

The image of this correspondence is extremely concrete: $X^*(T)$ is a lattice (that is, it is isomorphic to $\mathbb{Z}^\ell$), and W is a finite group of automorphisms of T (that is, of invertible $\ell \times \ell$ integer matrices). An extremal weight set is therefore an orbit of the finite group W on ℓ-tuples of integers.

Example 3.3. Suppose $K = O(2)$, the (disconnected) group of 2×2 real orthogonal matrices. The (unique) maximal torus in K is

$$T = SO(2) = \left\{ r(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}.$$

The group of characters of T is $X^*(T) = \mathbb{Z}$; the character $\ell \in \mathbb{Z}$ sends $r(\theta)$ to $\exp(i\ell\theta)$. The Weyl group $W = W(O(2), SO(2))$ is $\{\pm 1\}$. The Cartan-Weyl theorem therefore says that each irreducible representation of $O(2)$ has a unique pair of largest weights $\pm\ell$, and that there are just finitely many representations with extremal weights $\pm\ell$. In fact for $\ell = 0$ there are exactly two (one-dimensional) representations of $O(2)$ having 0 as the only weight (that is, being trivial on $SO(2)$). These are the trivial representation and the determinant representation. For $\ell \neq 0$, there is a unique (two-dimensional) representation of $O(2)$ having $\pm\ell$ as its only weights.

Example 3.4. Suppose $K = U(n)$. We may choose $T = U(1)^n$, so that $X^*(T) = \mathbb{Z}^n$: the character $\lambda = (\lambda_1, \ldots, \lambda_n)$ sends the matrix

$$d(\theta) = \begin{pmatrix} \exp(i\theta_1) & 0 & \cdots & 0 \\ 0 & \exp(i\theta_2) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \exp(i\theta_n) \end{pmatrix}$$

to $\exp\left(i \sum_j \lambda_j \theta_j\right)$. The Weyl group is the symmetric group Σ_n, acting by permutation of the coordinates. The standard inner product on $\mathbb{Z}^n$ is W-invariant.

A W-orbit on $X^*(T)$ may be identified with a nonincreasing sequence

$$(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n) \in \mathbb{Z}^n.$$

The Cartan-Weyl theorem says that every irreducible representation of $U(n)$ has a unique largest weight corresponding to such a nonincreasing sequence of integers; and that irreducible representations of $U(n)$ are in this way parametrized by nonincreasing sequences. The largest nonincreasing weight is called the *highest weight* of τ.

Here is an example. Suppose $E = S^k(\mathbb{C}^n)$, the space of homogeneous polynomials of degree k in n variables. This is a representation of $U(n)$ by the change of variables action. Each monomial

$$x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \sum a_j = k$$

is a weight vector for E, of weight

$$(a_1, a_2, \cdots, a_n) \in \mathbb{N}^n, \qquad \sum a_j = k.$$

The length of this weight is $\sum a_j^2$. We have

$$k^2 = \left(\sum a_j\right)^2 = \sum_j a_j^2 + 2\sum_{j<k} a_j a_k.$$

Every term in the last sum is nonnegative; so the conclusion is that

$$\sum_j a_j^2 \le k^2,$$

with equality if and only if just one of the a_j is nonzero. This means that the extremal weights of E are precisely the weights

$$(0, \dots, 0, k, 0, \dots, 0),$$

which do indeed form a single orbit of Σ_n. The highest weight of $S^k(\mathbb{C}^n)$ is therefore $(k, 0, \dots, 0)$.

For example, if $n = 2$ and $k = 3$, the representation $S^3(\mathbb{C}^2)$ has extremal weights $\{(3,0), (0,3)\}$; its other weights are $\{(2,1), (1,2)\}$.

Because we are interested in the size of representations, we want to understand the relationship between extremal weights and dimension.

Definition 3.5. The nonzero weights of the adjoint representation of K on $\mathfrak{k}_{\mathbb{C}}$ are called the *roots of T in K*:

$$R(K,T) = \{\alpha \in X^*(T)\backslash\{0\} \mid m_{\mathfrak{k}_{\mathbb{C}}}(\alpha) \neq 0\}.$$

We also need the corresponding *coroots* $R^\vee(K,T)$:

$$\alpha^\vee \in X_*(T) =_{\text{def}} \frac{1}{2\pi i} \ker \exp \subset i\mathfrak{t};$$

these are among other things integer-valued linear functionals on $X^*(T)$. The kernel of a coroot is therefore a codimension one sublattice of $X^*(T)$. Because the set of coroots is finite, these sublattices cannot cover all of $X^*(T)$. An element $\lambda \in X^*(T)$ is called *regular* if $\lambda(\alpha^\vee) \neq 0$ for all $\alpha \in R^\vee(K,T)$.

Each regular λ defines a *positive root system*

$$R^+(K,T) = \{\alpha \in R(K,T) \mid \alpha^\vee(\lambda) > 0\}.$$

Correspondingly, given R^+, we say that $\mu \in X^*(T)$ is *dominant for R^+* if

$$\mu(\alpha^\vee) \ge 0 \qquad (\alpha \in R^+).$$

Finally, we define

$$2\rho(R^+) = \sum_{\alpha \in R^+} \alpha \in X^*(T).$$

The weight $2\rho(R^+)$ turns out to be dominant and regular. Every coroot takes even values on $2\rho(R^+)$; nevertheless

$$\rho(R^+) =_{\mathrm{def}} \frac{1}{2} 2\rho(R^+) \in i\mathfrak{t}^*$$

may not belong to $X^*(T)$.

Theorem 3.6 (Weyl dimension formula)**.** *In the setting of Theorem 3.2, suppose that τ is an irreducible representation of K having an extremal weight $\mu \in X^*(T)$. Choose $R^+(K,T)$ so that μ is dominant (Definition 3.5. Then*

$$\dim(\tau) = \textit{(mult. of } \mu \textit{ in } \tau\textit{)} \cdot \textit{(number of dom wts in } W \cdot \mu\textit{)} \cdot \prod_{\alpha \in R^+} \frac{(\mu+\rho)(\alpha^\vee)}{\rho(\alpha^\vee)}.$$

The first two factors in the formula for $\dim(\tau)$ are small (bounded by $\#(K/K_0)$, for example). The last factor is a polynomial function of $\mu + \rho$, homogeneous of degree equal to half the dimension of K/T.

4 Representations of real reductive Lie groups

Assume now that

$$G \simeq K \times \mathfrak{s} \tag{4.1}$$

is a real reductive Lie group as in Definition 2.3. We begin with Harish-Chandra's basic results about representations of G.

Definition 4.2. A *representation* of G is a complete locally convex topological vector space V over $\mathbb{C}$, equipped with a homomorphism

$$\pi \colon G \to \mathrm{Aut}(V)$$

with the property that the map

$$G \times V \to V, \qquad (g, v) \mapsto \pi(g)v$$

is continuous. A *subrepresentation* is a closed subspace $U \subset V$ with the property that

$$\pi(G)U \subset U. \tag{4.2a}$$

We say that (π, V) is *irreducible* if there are exactly two subrepresentations (which must then be equal to 0 and V, with $V \neq 0$).

The set of *smooth vectors of* π is

$$V^\infty = \{v \in V \mid G \to V,\ g \mapsto \pi(g)v \text{ is a smooth map}\}. \tag{4.2b}$$

The space V^∞ has a natural structure of complete locally convex topological vector space (obtained by adding to the seminorms $|\cdot|_\alpha$ defining V some additional seminorms

$$|v|_{u,\alpha} =_{\text{def}} |u \cdot v|_\alpha$$

for u in the enveloping algebra $U(\mathfrak{g}_\mathbb{C})$). The representation π^∞ of G on V^∞ is again continuous, and it can be differentiated to get a Lie algebra representation of $\mathfrak{g}$ on V^∞. The smooth vectors are dense in V.

For each $(\tau, E) \in \widehat{K}$ we can define the *τ-isotypic subspace*

$$V_\tau = \text{largest sum of copies of } (E, \tau) \text{ in } (\pi, V) \text{ restricted to } K. \tag{4.2c}$$

The *multiplicity of τ in π* is

$$m_V(\tau) = m_\pi(\tau) = \dim \operatorname{Hom}_K(E, V). \tag{4.2d}$$

Then

$$\dim V_\tau = m_\pi(\tau) \cdot \dim \tau,$$

where of course both sides may be infinite. We say that π is *admissible* if all the multiplicities $m_\pi(\tau)$ are finite.

The set of *K-finite vectors*

$$V_K = \{v \in V \mid \dim\langle \pi(K)v \rangle < \infty\} \tag{4.2e}$$

clearly contains each V_τ; in fact

$$V_K = \bigoplus_{\tau \in \widehat{K}} V_\tau.$$

This subspace is dense in V; in fact

$$V_K^\infty = V_K \cap V^\infty \tag{4.2f}$$

is dense in V. This last space is preserved by the action of $\mathfrak{g}_\mathbb{C}$ on V^∞ and by the action of K on V_K; the representations π^∞ of $\mathfrak{g}_\mathbb{C}$ and π_K of K satisfy

$$\begin{aligned} \pi_K &= \text{locally finite representation of } K \\ d\pi_K &= \pi^\infty|_\mathfrak{k} \\ \pi_K(k)\pi^\infty(X) &= \pi^\infty(\operatorname{Ad}(k)X)\pi_K(k) \qquad (k \in K,\ X \in \mathfrak{g}). \end{aligned} \tag{4.2g}$$

Any simultaneous representation of the Lie algebra $\mathfrak{g}$ and the group K satisfying these three conditions is called a *$(\mathfrak{g}, K)$-module*. The $(\mathfrak{g}, K)$-module V_K^∞ is called the *Harish-Chandra module of π*. Two representations (π, V) and (π', V') are called *infinitesimally equivalent* if their Harish-Chandra modules are isomorphic.

Define

$$\mathfrak{Z}(\mathfrak{g}) = U(\mathfrak{g}_\mathbb{C})^G, \tag{4.2h}$$

the algebra of G-invariant elements of the enveloping algebra. (If G is connected, this is the center of the enveloping algebra, which is a polynomial ring. If G is disconnected, it is the invariants of a finite group acting on a polynomial ring.) We say that (π, V) is *quasisimple* if $\mathfrak{Z}(\mathfrak{g})$ acts by scalars on V^∞. If π is quasisimple, we write

$$\pi^\infty(z) = \chi_\pi(z)\,\mathrm{Id}_V, \qquad \chi_\pi \colon \mathfrak{Z}(\mathfrak{g}) \to \mathbb{C}; \tag{4.2i}$$

the algebra homomorphism χ_π is called the *infinitesimal character of* π.

We can clearly define the multiplicity function

$$m_X \colon \widehat{K} \to \mathbb{N} \cup \{\infty\}$$

for any $(\mathfrak{g}, K)$-module X. The terms *admissible* (all values of m_X are finite) and *quasisimple* ($\mathfrak{Z}(\mathfrak{g})$ acts by scalars) make sense for X as well.

Theorem 4.3 (Harish-Chandra)**.** *If V is an irreducible representation of a real reductive Lie group G, then the following conditions (explained in Definition 4.2) are equivalent:*

1. *V is quasisimple;*
2. *V is admissible;*
3. *the $(\mathfrak{g}, K)$-module V_K^∞ is irreducible.*

These conditions are automatically satisfied if V is an irreducible unitary *representation.*

Conversely, any irreducible $(\mathfrak{g}, K)$-module X is automatically quasisimple and admissible; and there is an irreducible representation V of G with $V_K^\infty \simeq X$.

The point of this theorem is that Schur's lemma (which need not hold for infinite-dimensional topological representations) suggests that irreducible representations *should* be quasisimple: that failure of $\mathfrak{Z}(\mathfrak{g}_\mathbb{C})$ to act by scalars is pathological. That irreducible unitary representations are automatically quasisimple is further evidence of this. There is no such obvious reason to expect admissibility to hold; this is a powerful consequence of the quasisimplicity hypothesis, which is perhaps the main point of Harish-Chandra's theorem. That admissibility implies (3) is fairly easy.

The fact that an irreducible $(\mathfrak{g}, K)$ module automatically "exponentiates" to an irreducible representation of G is closely related to the Cartan decomposition of Proposition 2.2. The construction of the representation can be thought of as solving differential equations on G with initial conditions along K; that G/K looks like $\mathbb{R}^N$ is the reason there are no global obstructions to this process.

We define

$$\begin{aligned} \widehat{G} &= \text{infinitesimal equivalence classes of} \\ &\qquad \text{irreducible quasisimple representations of } G \\ &= \text{equivalence classes of irreducible } (\mathfrak{g}, K)\text{-modules} \end{aligned} \tag{4.4}$$

Harish-Chandra's theorem says that we can attach to any $(\pi, V) \in \widehat{G}$ a multiplicity function

$$m_\pi \colon \widehat{K} \to \mathbb{N} \tag{4.5a}$$

According to the Cartan-Weyl Theorem 3.2, the domain of this function is (more or less) a lattice modulo a finite group; so it is something extremely concrete and combinatorial. For $(\pi, V) \in \widehat{G}$, we can define

$$V_N = \sum_{\tau \in \widehat{K},\ \|\lambda(\tau)\| \leq N} V_\tau. \tag{4.5b}$$

Here $\lambda(\tau) \in X^*(T)$ is an extremal weight (Theorem 3.2), which is well defined up to the action of W; and $\|\cdot\|$ is the norm from the inner product defined in (3.1o). The sum is therefore over a finite set of representations of K, so

$$f_V(N) =_{\mathrm{def}} \dim V_N = \sum_{\tau \in \widehat{K},\ \|\lambda(\tau)\| \leq N} m_V(\tau) \dim \tau \tag{4.5c}$$

is a weakly increasing $\mathbb{N}$-valued function. According to the Weyl dimension formula Theorem 3.6, $\dim \tau$ depends more or less in a polynomial way on the extremal weight $\lambda(\tau)$. The following proposition says that the multiplicities $m_V(\tau)$ also depend in a reasonable way on $\lambda(\tau)$.

Proposition 4.6. *(see [2]) There is a nonnegative integer $d = d(\pi)$ and a positive constant $c = c(\pi)$ so that*

$$f_\pi(N) \sim c(\pi) N^{d(\pi)}$$

as $N \to \infty$ (in the sense that the ratio converges to 1). The integer $d(\pi)$ is equal to half the real dimension of some nilpotent orbit of G on $\mathfrak{g}^*$.

The nonnegative integer $d(\pi)$ is called the *Gelfand-Kirillov dimension of* π. In Section 5 we will try to understand how to interpret $d(\pi)$.

5 Compact homogeneous spaces

In this section we will look at some examples of irreducible representations of real reductive groups, calculate their Gelfand-Kirillov dimensions, and try in this way to get some understanding of the meaning of this invariant. We will work with

$$G = GL(n, \mathbb{R}), \tag{5.1a}$$

the group of invertible linear transformations of $\mathbb{R}^n$. We fix a collection of positive integers

$$\mathbf{d} = (d_1, d_2, \cdots, d_k), \qquad \sum_{i=1}^k d_i = n. \tag{5.1b}$$

A *flag of type* **d** is a chain of vector subspaces

$$\mathcal{E} = (0 = E_0 \subset E_1 \subset \cdots \subset E_k = \mathbb{R}^n)\,, \qquad \dim E_i/E_{i-1} = d_i. \tag{5.1c}$$

The *partial flag manifold* is

$$X(\mathbf{d}) = \text{all flags of type } \mathbf{d} \text{ in } \mathbb{R}^n. \tag{5.1d}$$

The fundamental theorems of linear algebra assert that $X(\mathbf{d})$ is a homogeneous space for $G = GL(n, \mathbb{R})$. There is a natural base point

$$\mathcal{E}_0 = \left(0 = \mathbb{R}^0 \subset \mathbb{R}^{d_1} \subset \mathbb{R}^{d_1+d_2} \subset \cdots \subset \mathbb{R}^{d_1+\cdots+d_k} = \mathbb{R}^n\right). \tag{5.1e}$$

The stabilizer of $\mathcal{E}_0$ is the group of block upper triangular matrices

$$P(\mathbf{d}) = \left\{ \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & \cdots & A_k \end{pmatrix} \mid A_i \in GL(d_i, \mathbb{R}) \right\}, \tag{5.1f}$$

which is called a *standard parabolic subgroup of type* **d**. Here the first asterisk represents an arbitrary matrix of size $d_1 \times d_2$, and so on. More generally, if $g \in GL(n, \mathbb{R})$ is an arbitrary matrix with columns $(C_1, \ldots, C_n)$, then

$$g \cdot \mathcal{E}_0 = (0 \subset \langle C_1, \ldots, C_{d_1} \rangle \subset \langle C_1, \ldots, C_{d_1+d_2} \rangle \subset \cdots \subset \langle C_1, \ldots, C_n \rangle = \mathbb{R}^n)\,.$$

Similarly, we can define an *orthogonal flag of type* **d** to be a collection of mutually orthogonal vector subspaces

$$\mathcal{F} = (F_1, \ldots, F_k), \qquad F_i \perp F_j, \qquad \dim F_i = d_i. \tag{5.1g}$$

The *orthogonal flag manifold* $Y(\mathbf{d})$ is the collection of all orthogonal flags of type **d**. Just as for flags above, we find that $Y(\mathbf{d})$ is a homogeneous space for $O(n)$; the isotropy group of the natural base point is

$$O(\mathbf{d}) =_{\text{def}} O(d_1) \times O(d_2) \times \cdots \times O(d_k) = O(n) \cap P(\mathbf{d}). \tag{5.1h}$$

The maps

$$F_i = \text{orthogonal complement of } E_{i-1} \text{ in } E_i$$
$$E_i = \langle F_1, F_2, \cdots, F_i \rangle$$

are mutually inverse $O(n)$-equivariant bijections between $X(\mathbf{d})$ and $Y(\mathbf{d})$. Consequently

$$G/P(\mathbf{d}) \simeq X(\mathbf{d}) \simeq Y(\mathbf{d}) \simeq O(n)/O(\mathbf{d}) \tag{5.1i}$$

The reason for singling out the homogeneous spaces $X(\mathbf{d})$ is that they arise as "limits" of more or less arbitrary homogeneous spaces. Here is a concrete result in that direction.

Proposition 5.2. *Suppose $\mathfrak{h} \subset \mathfrak{gl}(n, \mathbb{R})$ is any Lie subalgebra. Then there is a Lie subalgebra $\mathfrak{h}^1$ and a standard parabolic subgroup $P(\mathbf{d})$ with the following properties.*

1. *The Lie algebra $\mathfrak{h}^1$ is a limit of conjugates of $\mathfrak{h}$. That is, there is a sequence (g_m) of elements of G so that*

$$\lim_{m\to\infty} \mathrm{Ad}(g_m)(\mathfrak{h}) = \mathfrak{h}^1.$$

   The limit is taken in the Grassmann variety of $\dim \mathfrak{h}$*-dimensional linear subspaces of $\mathfrak{g}$.*

2. *The Lie algebra $\mathfrak{h}^1$ is an ideal in $\mathfrak{p}(\mathbf{d})$. More precisely, $\mathfrak{p}(\mathbf{d})$ is equal to the normalizer in $\mathfrak{g}$ of $\mathfrak{h}^1$.*

Proof. Write $m = \dim \mathfrak{h}$, and $Z(m)$ for the (compact) Grassmann variety of m-dimensional subspaces of $\mathfrak{g}$. The group G acts algebraically on $Z(m)$ by conjugation of subspaces. The orbit closure $\overline{\mathrm{Ad}(G)(\mathfrak{h})} \subset Z(m)$ is compact, and it is a union of G orbits, all except $Ad(G)(\mathfrak{h})$ being of strictly smaller dimension. (The algebraic nature of the action prevents more pathological behavior.) Being closed under commutator of matrices is a closed condition on subspaces of $\mathfrak{g}$, so everything in the closure is a Lie algebra. Any G-orbit $\mathrm{Ad}(G)(\mathfrak{h}^1)$ of minimal dimension in the closure must be compact: so

$$\mathrm{Ad}(G)(\mathfrak{h}^1) = G/N_G(\mathfrak{h}^1)$$

is a compact algebraic homogeneous space for G. Any such homogeneous space is isomorphic to $X(\mathbf{d})$. (For algebraic groups over an algebraically closed field, the fact that a projective homogeneous space must be G modulo a parabolic is standard. We omit the details of passage to the real group $GL(n, \mathbb{R})$.) □

The ideals in $\mathfrak{p}(\mathbf{d})$ are few and not difficult to understand; so the limiting homogeneous spaces arising from this proposition are closely related to $X(d)$.

Example 5.3. Suppose $H = O(n) \subset GL(n, \mathbb{R})$, so that

$$\begin{aligned}\mathfrak{h} &= n \times n \text{ real skew-symmetric matrices}\\ &= \{A = (a_{ij}) \mid a_{ij} = -a_{ji}\}.\end{aligned}$$

Conjugating by the diagonal matrix $g(\mathbf{r})$ with entries $\mathbf{r} = (r_1, \ldots, r_n)$ gives

$$\mathrm{Ad}(g(\mathbf{r}))(\mathfrak{h}) = \mathfrak{h}(\mathbf{r}) = \left\{A \mid r_i r_j^{-1} a_{ij} = -r_j r_i^{-1} a_{ji}\right\}.$$

If the positive real numbers r_i are strictly decreasing, then

$$\lim_{m\to\infty} \mathrm{Ad}(g(\mathbf{r})^m)(\mathfrak{h}) = \mathfrak{n} = \text{strictly upper triangular matrices.}$$

The normalizer of $\mathfrak{n}$ is the Borel subgroup

$$B = P((1, \ldots, 1)) = \text{upper triangular invertible matrices.}$$

The homogeneous space G/H can be identified with all positive definite quadratic forms on $\mathbb{R}^n$. If we fix a complete flag $\mathcal{E} = (E_0 \subset \cdots \subset E_n)$, then "going to infinity in the direction of $\mathcal{E}$" on G/H means looking at quadratic forms that are very large on the line E_1, fairly large on the plane E_2, and so on. Eigenfunctions of invariant differential operators on G/H can be studied by looking at their boundary values—appropriate limits at infinity—as sections of line bundles on $X(1,\ldots,1)$. A beautiful account may be found in [5].

Having established that the partial flag varieties $X(\mathbf{d})$ are interesting spaces to look at, let us use them to make some representations. If

$$X = G/H \tag{5.4a}$$

is a homogeneous space for a Lie group G, then an equivariant vector bundle $\mathcal{F}$ over X is exactly the same thing as a finite-dimensional representation (τ, F) of H. Given the vector bundle, F is just the fiber over the base point eH. Given the representation, we can recover the vector bundle as

$$\mathcal{F} = G \times_H F =_{\text{def}} G \times F/\sim, \qquad (gh, f) \sim (g, \tau(h)f). \tag{5.4b}$$

The space of smooth sections of $\mathcal{F}$ may be identified with

$$C^\infty(X, \mathcal{F}) \simeq \{\phi\colon G \to F,\ \phi(xh) = \tau(h)^{-1}\phi(x)\} \tag{5.4c}$$

Clearly the group G acts on $C^\infty(X, \mathcal{F})$ by left translation

$$[T(g)\phi](x) = \phi(g^{-1}x) \qquad (\phi \in C^\infty(X, \mathcal{F}), g \in G)\,. \tag{5.4d}$$

This is the (smooth) induced representation:

$$(T, C^\infty(X, \mathcal{F})) =_{\text{def}} \operatorname{Ind}_H^G(\tau, F). \tag{5.4e}$$

Suppose now that we are in the setting (5.1). Let us fix k finite-dimensional representations

$$(\tau_i, F_i) \in \widehat{GL(d_i, \mathbb{R})}, \qquad (i = 1, \ldots, k) \tag{5.4f}$$

so that

$$(\tau, F), \qquad F = F_1 \otimes \cdots \otimes F_k \in \widehat{P(\mathbf{d})} \tag{5.4g}$$

is an irreducible representation. (The various asterisk elements in (5.1f) all act trivially.) We therefore get a representation of $G = GL(n, \mathbb{R})$

$$(\operatorname{Ind}_{P(\mathbf{d})}^G(\tau), C^\infty(X(\mathbf{d}), \mathcal{F})) \tag{5.4h}$$

on the space of sections of a vector bundle over the flag variety $X(\mathbf{d})$. These are fairly typical representations of $GL(n, \mathbb{R})$. We would therefore like to understand what Gelfand-Kirillov dimension means for them.

Proposition 5.5. *Suppose we are in the setting* (5.4)*, so that* $\mathcal{F} \to X(\mathbf{d})$ *is a finite-dimensional equivariant vector bundle on the partial flag variety. Write*

$$(\pi, V) = \left(\mathrm{Ind}_{P(\mathbf{d})}^{GL(n,\mathbb{R})}(\tau), C^\infty(X(\mathbf{d}), \mathcal{F})\right)$$

for the corresponding induced representation.

1. *The representation* π *is a quasisimple representation of* $GL(n,\mathbb{R})$*, having a finite number of irreducible composition factors.*
2. *There is an* $O(n)$*-invariant Laplace operator* Δ *on* $C^\infty(X(\mathbf{d}), \mathcal{F})$*, given by the action of the Casimir operator* Ω_K *of* $O(n)$*. The growth function* f_V *of* (4.5c) *is approximately the eigenvalue distribution function for* Δ*:*

   $$f_V(N) \approx \textit{number of eigenvalues less than or equal to } N^2.$$

3. *The asymptotic behavior of* f_V *is*

   $$f_V(N) \sim c(X(\mathbf{d}))(\dim F)N^{\dim X(\mathbf{d})}.$$

   That is, the Gelfand-Kirillov dimension of π *is equal to the dimension of the partial flag variety* $X(\mathbf{d})$*:*

   $$d(\pi) = \dim X(\mathbf{d}) = [n^2 - \sum d_i^2]/2 = \sum_{i<j} d_i d_j.$$

4. *The multiplicity function* m_V *may be computed by Frobenius reciprocity: for any irreducible representation* (τ, E) *of* $O(n)$*,*

   $$m_V(\tau) = \dim \mathrm{Hom}_{O(d_1)\times\cdots\times O(d_k)}\left(\tau, (F_1|_{O(d_1)}) \otimes \cdots \otimes (F_k|_{O(d_k)})\right).$$

Proof. That the composition series is finite is a theorem of Harish-Chandra. Since our invariant inner product

$$\langle X, Y\rangle = \mathrm{tr}(XY)$$

is negative definite on $\mathfrak{k}$, the natural Laplace operator is *minus* a sum of squares. The operator Δ therefore has *nonnegative* eigenvalues. The Casimir operator acts by a scalar $\tau(\Omega_K)$ on each $\tau \in \widehat{K}$, so the eigenvalue distribution function is

$$\text{multiplicity of eigenvalue } \mu = \sum_{\tau\in\widehat{K}, \tau(\Omega_K)=\mu} \dim V_\tau.$$

If τ has dominant extremal weight λ, then

$$\tau(\Omega_K) = \langle \lambda + 2\rho, \lambda\rangle = \langle\lambda,\lambda\rangle + \text{error of size } \langle\lambda,\lambda\rangle^{1/2}.$$

That is, the size of the eigenvalue is approximately the length squared of λ. It follows that $f_V(N)$ and the number of eigenvalues up to N^2 have the same asymptotic behavior. The formula in (3) for the eigenvalues is Weyl's law: the constant $c(X(\mathbf{d}))$ involves things like the volume of $X(\mathbf{d})$.

For the last statement, (5.1i) guarantees that

$$\pi|_K \simeq \mathrm{Ind}_{O(d_1)\times\cdots\times O(d_k)}^{O(n)} \left((F_1|_{O(d_1)}) \otimes \cdots \otimes (F_k|_{O(d_k)})\right).$$

Therefore Frobenius reciprocity for $K = O(n)$ computes $\mathrm{Hom}_K(\tau, \pi|_K)$, giving the answer stated. □

The conclusion is that lots of interesting representations of $GL(n, \mathbb{R})$ appear in sections of vector bundles on partial flag varieties. For such representations, *the Gelfand-Kirillov dimension is the dimension of the underlying partial flag variety.*

6 The wavefront set of a representation

In this section and the next we recall two geometric descriptions of the Gelfand-Kirillov dimension. We suppose throughout this section that

$$(\pi, V) \in \widehat{G} \tag{6.1}$$

is an irreducible quasisimple representation of the real reductive Lie group G. If ϕ is a test density on G (that is, a smooth measure of compact support) then

$$\pi(\phi) = \int_G \pi(g)\phi(g) \in \mathrm{End}(V) \tag{6.2a}$$

is a well-defined linear map. Harish-Chandra has shown that $\pi(\phi)$ has a well-defined trace, and that

$$\Theta_\pi(\phi) =_{\mathrm{def}} \mathrm{tr}(\pi(\phi)) \tag{6.2b}$$

is a generalized function on G, which depends only on the infinitesimal equivalence class of π. We compute easily that

$$\pi(\mathrm{Ad}(g)\phi) = \pi(g)\pi(\phi)\pi(g)^{-1}. \tag{6.2c}$$

Because the trace of a linear operator is unchanged by conjugation, it follows that

$$\mathrm{Ad}(g)(\Theta_\pi) = \Theta_\pi \tag{6.2d}$$

If $u \in U(\mathfrak{g}_{\mathbb{C}})$ acts on generalized functions by differentiation on the left, then

$$\pi(u \cdot \phi) = \pi^\infty(u)\pi(\phi); \tag{6.2e}$$

the equation makes sense because $\pi(\phi)$ is easily seen to carry V into V^∞. (This fact is part of the proof that V^∞ is dense in V.)

If $z \in \mathfrak{Z}(\mathfrak{g})$ is central, (6.2e) amounts to $\pi(z \cdot \phi) = \chi_\pi(z)\pi(\phi)$. Therefore

$$z \cdot \Theta_\pi = \chi_\pi(z)\Theta_\pi. \tag{6.2f}$$

The two properties (6.2d) and (6.2f) are expressed by saying that Θ_π is an *invariant eigendistribution*. (This terminology was fixed before the importance of distinguishing generalized functions (linear functionals on test densities) from distributions (linear functionals on test functions) was so clear.)

Any distribution on a manifold M has a well-defined *wavefront set*, which is a closed cone in the cotangent bundle $T^*(M)$. Howe in [6] defines the *wavefront set of* π to be

$$\mathrm{WF}(\pi) = \mathrm{WF}(\Theta_\pi) \cap (T_e^*(G) = \mathfrak{g}^*)\,. \tag{6.2g}$$

It is a consequence of (6.2d) and (6.2f) (because of (6.4c) below) that

$$\mathrm{WF}(\pi) = \text{ closed union of nilpotent orbits of } \mathrm{Ad}(G) \text{ on } \mathfrak{g}^* \tag{6.2h}$$

Proposition 6.3 ([2])**.** *In the setting* (6.2)*, the Gelfand-Kirillov dimension of* π *is half the (real) dimension of* $\mathrm{WF}(\pi)$*.*

Notice that the dimension of $\mathrm{WF}(\pi)$ is well-defined because the set is a finite union of homogeneous spaces for the Lie group G.

To go further, it is useful to have in hand some basic structure theory about reductive Lie algebras.

Definition 6.4. Suppose $G = K \times \mathfrak{s}$ (Definition 2.3) is the Cartan decomposition of our real reductive Lie group, and

$$\mathfrak{g} = \mathfrak{k} + \mathfrak{s} \tag{6.4a}$$

is the Cartan decomposition of its Lie algebra. Write $\mathfrak{g}_{\mathbb{C}}$ for the complexification of $\mathfrak{g}$, and

$$K_{\mathbb{C}} = \text{complexification of } K, \tag{6.4b}$$

a complex reductive algebraic group.

The *complex-nilpotent cone* is

$$\mathcal{N}_{\mathbb{C}}^* = \left\{\lambda \in \mathfrak{g}_{\mathbb{C}}^* \mid p(\lambda) = 0, \quad p \in (\mathfrak{g}S(\mathfrak{g}))^G\right\}. \tag{6.4c}$$

(The requirement is that every $\mathrm{Ad}(G)$-invariant polynomial of strictly positive degree should vanish at λ. This description of nilpotent elements is classical linear algebra in the case of $GL(n)$: an $n \times n$ nilpotent matrix λ is one whose characteristic polynomial is x^n; and the coefficient of x^{n-d} in the characteristic polynomial of any matrix λ is the value at λ of conjugation-invariant polynomial of degree d. For general reductive groups this description of nilpotent elements is due to Kostant.)

The *real-nilpotent cone* is

$$\mathcal{N}_{\mathbb{R}}^* = \mathcal{N}_{\mathbb{C}}^* \cap \mathfrak{g}, \tag{6.4d}$$

the nilpotent linear functionals taking real values on the real Lie algebra. We saw in (6.2h) that

$$\mathrm{WF}(\pi) \subset \mathcal{N}_{\mathbb{R}}^*. \tag{6.4e}$$

The θ-*nilpotent cone* is

$$\mathcal{N}_\theta^* = \mathcal{N}_\mathbb{C}^* \cap \mathfrak{s}_\mathbb{C}^*. \tag{6.4f}$$

We will see in (7.10d) below that

$$\mathrm{AV}(\pi) \subset \mathcal{N}_\theta^*. \tag{6.4g}$$

Theorem 6.5 (Kostant-Sekiguchi [15]). *Suppose that the real reductive Lie group G admits a complexification $G_\mathbb{C}$.*

1. *$G_\mathbb{C}$ acts on $\mathcal{N}_\mathbb{C}^*$ (*cf. (6.4c)*) with finitely many orbits, each a complex symplectic manifold and therefore of even complex dimension.*
2. *G acts on $\mathcal{N}_\mathbb{R}^*$ (*cf. (6.4d)*) with finitely many orbits, each a real form of the corresponding complex orbit and therefore of even real dimension.*
3. *$K_\mathbb{C}$ acts on $\mathcal{N}_\theta^*$ (*cf. (6.4f)*) with finitely many orbits, each a complex Lagrangian submanifold of the corresponding complex orbit.*

The reader may dislike the lack of parallel between the "real form" in (2) and the "Lagrangian" in (3). If the complex symplectic form is multiplied by i, then the real form in (2) and the complex Lagrangian in (3) both become Lagrangian submanifolds of the underlying real symplectic manifold in (1). This is a useful perspective for trying to construct representations from coadjoint orbits, but we will not need it here.

The wavefront set $\mathrm{WF}(\pi)$ is a geometric invariant constructed by (soft) real analysis. In Section 7 we will consider a geometric invariant $\mathrm{AV}(V_K^\infty)$ constructed by (soft) algebraic geometry. Here is a result relating the two invariants to each other and to Gelfand-Kirillov dimension.

Theorem 6.6 (Schmid-Vilonen [14]). *For any irreducible quasisimple representation (π, V) of the real reductive Lie group G, the wavefront set $\mathrm{WF}(\pi)$ corresponds precisely to the associated variety $\mathrm{AV}(V_K^\infty)$ under the Kostant-Sekiguchi correspondence (Theorem 6.5). In particular,*

$$\dim_\mathbb{R}(\mathrm{WF}(\pi)) = 2\dim_\mathbb{C}\left(\mathrm{AV}(V_K^\infty)\right).$$

Consequently

$$\textit{Gelfand-Kirillov dimension of } \pi = \dim_\mathbb{C}\left(\mathrm{AV}(V_K^\infty)\right).$$

The difficult part of this theorem is the first statement. The first displayed formula is then a consequence of the statements in Theorem 6.5, and the second of Proposition 6.3.

We will view this theorem as making Gelfand-Kirillov dimension a very crude version of the much more subtle and interesting geometric invariants $\mathrm{WF}(\pi)$ and $\mathrm{AV}(V_K^\infty)$. In sections 7 and 8 we will describe a further refinement of this invariant; and finally in section 9 we will describe an algorithm for computing the refined invariant (and therefore also computing associated varieties and wave-front sets). As often happens in geometric representation theory, the algorithm for computing the refined invariant cannot be simplified to compute associated varieties directly: the algorithm depends in an essential way on the refined information.

7 Understanding associated varieties

We have motivated Gelfand-Kirillov dimension (for a representation of G) in terms of the (real) dimensions of certain compact homogeneous spaces for G (on which such representations sometimes live). Theorem 6.6 relates Gelfand-Kirillov dimension to the (complex) dimensions of certain algebraic homogeneous spaces—no longer for G, but for a closely related complex reductive algebraic group $K(\mathbb{C})$. In this section we will outline a more direct path from representations to these algebraic homogeneous spaces.

It is an idea dating back to the creation of quantum mechanics that a group representation is in some sense a "quantum mechanical" object. I will not try to justify that idea in great detail (although I will commend to the reader's attention Mackey's account in [13]). For us it is enough to note that unitary group representations consist of operators on Hilbert spaces, just as the mathematical models of quantum mechanics are usually phrased in terms of operators on Hilbert spaces.

A fundamental idea in quantum physics is that of "classical limit:" that a quantum-mechanical system can be thought of as depending on a small real parameter h (Planck's constant); and that in the limit as $h \to 0$, the behavior of the system approximates that of a corresponding classical-mechanical system. Kirillov and Kostant were among the first to bring that idea into representation theory. They formulated "classical" analogues of group representations, studied those classical analogues, and then tried to prove properties of the group representations analogous to what they found in the classical picture.

Our goal here is similar in spirit, although the precise technical connections are tenuous. Because we will be doing *complex* algebraic geometry, what is central is the *complexified* Lie algebra

$$\mathfrak{g}_{\mathbb{C}} = (\mathrm{Lie}(G) \otimes_{\mathbb{R}} \mathbb{C}) \tag{7.7a}$$

We will think of a representation of a Lie group G (always on a *complex* vector space) as being first of all a *module* for the complex universal enveloping algebra

$$U(\mathfrak{g}_{\mathbb{C}}) =_{\mathrm{def}} (\text{tensor algebra of } \mathfrak{g}_{\mathbb{C}}) / \langle x \otimes y - y \otimes x - [x, y] \mid x, y \in \mathfrak{g}_{\mathbb{C}} \rangle; \tag{7.7b}$$

and second as an *algebraic representation* for the complex reductive algebraic group $K_{\mathbb{C}}$. If we are starting with any Lie group representation V, this $U(\mathfrak{g}_{\mathbb{C}})$-module could be V^∞. If (π, V) is a quasisimple irreducible representation of a reductive Lie group, then we will usually want to consider the Harish-Chandra $U(\mathfrak{g})$-module $V^\infty_{K_{\mathbb{C}}}$; an important technical feature is that

$$V^\infty_{K_{\mathbb{C}}} \text{ is a } \textit{finitely generated } U(\mathfrak{g}_{\mathbb{C}})\text{-module}, \tag{7.7c}$$

as follows easily from Harish-Chandra's Theorem 4.3.

For us the quantum-mechanical character of representations is captured by the *non-commutativity* of $U(\mathfrak{g}_{\mathbb{C}})$. To take classical limit, we wish to replace $U(\mathfrak{g}_{\mathbb{C}})$ by a commutative algebra. Here is the (classical) way to do that.

Theorem 7.8 (Poincare-Birkhoff-Witt). *If $\mathfrak{g}_{\mathbb{C}}$ is a complex Lie algebra, define*

$$U_n(\mathfrak{g}_{\mathbb{C}}) = \text{span of } \{x_1 \cdots x_m | x_i \in \mathfrak{g}_{\mathbb{C}},\ m \leq n\} \subset U(\mathfrak{g}_{\mathbb{C}}).$$

Then $\{U_n(\mathfrak{g}_{\mathbb{C}})\}$ is an increasing exhaustive algebra filtration, meaning for example that

$$U_p(\mathfrak{g}_{\mathbb{C}}) \cdot U_q(\mathfrak{g}_{\mathbb{C}}) \subset U_{p+q}(\mathfrak{g}_{\mathbb{C}}).$$

The associated graded algebra is the symmetric algebra of the vector space $\mathfrak{g}_{\mathbb{C}}$:

$$U_n(\mathfrak{g}_{\mathbb{C}})/U_{n-1}(\mathfrak{g}_{\mathbb{C}}) \simeq S^n(\mathfrak{g}_{\mathbb{C}}),$$

homogeneous polynomials of degree n.

The adjoint action Ad *of $K_{\mathbb{C}}$ on the enveloping algebra preserves $U_n(\mathfrak{g}_{\mathbb{C}})$, and descends to the natural action* Ad *on the symmetric algebra.*

Corollary 7.9. *Suppose V is a finitely-generated $(U(\mathfrak{g}_{\mathbb{C}}), K_{\mathbb{C}})$-module, with finite-dimensional $K_{\mathbb{C}}$-invariant generating subspace V_0. Define*

$$V_n = U_n(\mathfrak{g}_{\mathbb{C}}) \cdot V_0.$$

1. *The subspaces V_n are an increasing exhaustive filtration of V by finite-dimensional $K_{\mathbb{C}}$-invariant subspaces, and*

$$U_n(\mathfrak{g}_{\mathbb{C}}) \cdot V_m \subset V_{n+m}.$$

2. *The associated graded vector space*

$$\operatorname{gr} V = \bigoplus_{n=0}^{\infty} V_n/V_{n-1}$$

is in a natural way a finitely-generated graded $S(\mathfrak{g}_{\mathbb{C}}/\mathfrak{k}_{\mathbb{C}})$-module, carrying a compatible algebraic action of $K_{\mathbb{C}}$.

3. *Suppose in addition that V is quasisimple (see (4.2i)) so that V_0 is preserved by the action of $\mathfrak{Z}(\mathfrak{g}_{\mathbb{C}})$. Then*

$$(\mathfrak{g}_{\mathbb{C}} S(\mathfrak{g}_{\mathbb{C}}))^G \subset \operatorname{Ann}(\operatorname{gr} V).$$

Definition 7.10. Suppose E is a finite-dimensional complex vector space, and $\mathbb{C}[E]$ is the polynomial ring of algebraic functions on E. Any finitely generated $\mathbb{C}[E]$-module M has a *support*

$$\operatorname{Supp}(M) = \{\lambda \in E \mid f(\lambda) = 0 \quad (f \in \operatorname{Ann}(M))\}. \tag{7.10a}$$

This is a subvariety of E. If M is graded, then $\mathrm{Supp}(M)$ is a cone (that is, it is preserved by dilations). More generally, if the algebraic group H acts on E and M carries a compatible algebraic action of H, then $\mathrm{Supp}(M)$ is a union of H-orbits on E.

Suppose now that V is a finitely-generated $(U(\mathfrak{g}_{\mathbb{C}}), K_{\mathbb{C}})$-module. The *associated variety of* V is by definition

$$\mathrm{AV}(V) = \mathrm{Supp}(\mathrm{gr}(V)) \subset \mathfrak{g}_{\mathbb{C}}^*; \tag{7.10b}$$

here $\mathrm{gr}\, V$ is constructed as in Corollary 7.9. Because $\mathrm{gr}\, V$ is graded, $\mathrm{AV}(V)$ is a cone. We have

$$\mathfrak{k}_{\mathbb{C}} \subset \mathrm{Ann}(\mathrm{gr}\, V),$$

so by the definition of support,

$$\mathrm{AV}(V) \subset (\mathfrak{g}_{\mathbb{C}}/\mathfrak{k}_{\mathbb{C}})^*. \tag{7.10c}$$

If in addition V is quasisimple, condition (3) of Corollary 7.9 says that

$$(\mathfrak{g}_{\mathbb{C}} S(\mathfrak{g}_{\mathbb{C}}))^G \subset \mathrm{Ann}(\mathrm{gr}\, V),$$

and therefore (by Definition 6.4)

$$\mathrm{AV}(V) \subset \mathcal{N}_\theta^*, \tag{7.10d}$$

a Zariski-closed union of $K_{\mathbb{C}}$-orbits.

This definition is the "direct path" from representations of G to orbits of $K_{\mathbb{C}}$ on $\mathcal{N}_\theta^*$. To make it effective, one needs a way to control $\mathrm{Ann}(\mathrm{gr}\, V)$ for a $(U(\mathfrak{g}_{\mathbb{C}}), K_{\mathbb{C}})$-module V. This annihilator is spanned by homogeneous polynomials $p \in S^n(\mathfrak{g}_{\mathbb{C}})$ with the property that a representative $u \in U_n(\mathfrak{g}_{\mathbb{C}})$ for p (see Theorem 7.8) satisfies

$$u \cdot V_m \subset V_{m+n-1};$$

that is, that u does not raise the level of the filtration of V as much as expected. If we can find many such elements p, then we can prove that $\mathrm{AV}(V)$ is small. Corollary 7.9 shows that we can take for p anything in $\mathfrak{k}_{\mathbb{C}}$ or in $(\mathfrak{g}_{\mathbb{C}} S(\mathfrak{g}_{\mathbb{C}}))^G$; and from this we deduced that $\mathrm{AV}(V) \subset \mathcal{N}_\theta^*$. The first of these possibilities is "K-finite" and the second is "quasisimple:" the fundamental defining characteristics of Harish-Chandra modules. Finding more elements $p \in \mathrm{Ann}(\mathrm{gr}\, V)$ is very difficult, even when we have a very precise description of V.

In order to give an algorithm for computing $\mathrm{AV}(V)$, we will therefore proceed indirectly. The strategy is to *refine* the invariant $\mathrm{AV}(V)$: instead of just the variety $\mathrm{AV}(V)$ we want something like an equivariant vector bundle over this variety.

Definition 7.11. Suppose X is a complex algebraic variety and suppose that H is a complex algebraic group acting on X. We are interested in the abelian category

$$\mathrm{Coh}^H(X) \tag{7.11a}$$

of H-equivariant coherent sheaves on X. The *equivariant K-theory of X* is the Grothendieck group $K^H(X)$ of $\mathrm{Coh}^H(X)$: the free abelian group generated by equivariant coherent sheaves, subject to the relations

$$[A] + [C] = [B] \quad \text{whenever} \quad 0 \to A \to B \to C \to 0 \tag{7.11b}$$

is a short exact sequence in $\mathrm{Coh}^H(X)$.

A nice general reference for equivariant K-theory, covering most of what we will need, is [3, Chapter 5]. The last word on all issues is [16].

We are studying the category

$$\mathcal{M}(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})\text{–}\mathrm{Mod}^{\mathrm{fin}} \tag{7.12a}$$

of finite length $(\mathfrak{g}, K)$-modules. The assumption of finite length is equivalent to the assumption of both finite generation (which we need to have a nice theory of associated graded modules) and annihilation by an ideal of finite codimension in $\mathfrak{Z}(\mathfrak{g})$ (which we need to make associated varieties nilpotent; see (6.4c)). We write

$$\begin{aligned} K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}}) &= \text{Grothendieck group of } \mathcal{M}(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})\text{–}\mathrm{Mod}^{\mathrm{fin}} \\ &= \text{free } \mathbb{Z}\text{-module with basis } \widehat{G}. \end{aligned} \tag{7.12b}$$

(*cf.* (4.4)). The map

$$(V + (\text{choice of good filtration})) \mapsto \operatorname{gr} V \tag{7.12c}$$

(explained in Corollary 7.9 and Definition 7.10) descends to a homomorphism of Grothendieck groups

$$[\mathrm{gr}] \colon K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}}) \to K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*) \tag{7.12d}$$

which is *independent* of the choice of good filtration used to construct it. (A proof of this independence may be found in [18].) We will see that this map to K-theory is a refinement of the associated variety construction. In order to do that, we need a more information about both the domain and the range of the map.

On the domain side, one issue is to understand the kernel of [gr]. For that, we make use of the category

$$\mathcal{M}(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})\text{–}\mathrm{Mod}^{\mathrm{fin}}{}_{K_{\mathbb{C}}} : \tag{7.12e}$$

the objects are still finite length $(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})$-modules, but now for morphisms we allow any linear map respecting the action of $K_{\mathbb{C}}$. Because there are more morphisms, and so more short exact sequences, the Grothendieck group is smaller:

$$\begin{aligned} &K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})_{K_{\mathbb{C}}} = \text{Groth. gp. of } (\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})\text{-modules restr. to } K_{\mathbb{C}}, \\ &\quad K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}}) \twoheadrightarrow K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})_{K_{\mathbb{C}}}. \end{aligned} \tag{7.12f}$$

If we write

$$\widehat{G}_{\text{temp},\mathbb{R}} \tag{7.12g}$$

for the (discrete) set of tempered irreducible representations with real infinitesimal character, then

$$K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})_{K_{\mathbb{C}}} = \text{free } \mathbb{Z}\text{-module with basis } \widehat{G}_{\text{temp},\mathbb{R}}. \tag{7.12h}$$

(This is essentially proven in [19].)

On the domain side, we will also make use of the Langlands classification. Suppose $P = MAN$ is a parabolic subgroup of G, and $\delta \in \widehat{M}$ is a tempered representation of M. A character $\nu \in \widehat{A} \simeq \mathfrak{a}_{\mathbb{C}}^*$ is called *strictly dominant* if

$$\operatorname{Re}\langle \nu, \alpha \rangle > 0, \qquad \text{(every weight } \alpha \text{ of } A \text{ in } \mathfrak{n}). \tag{7.12i}$$

If ν is strictly dominant, then the representation

$$I(\delta \otimes \nu) = \operatorname{Ind}_{MAN}^{G}(\delta \otimes \nu \otimes 1) \tag{7.12j}$$

is called a *standard representation of* G.

The Langlands classification theorem says that standard representations and irreducible representations are related in the same way as Verma modules and irreducible highest weight modules.

Theorem 7.13 (Langlands [11])**.** *Suppose G is a real reductive Lie group. Use the notation just described.*

1. *Each standard represenation $I(\delta \otimes \nu)$ is a quasisimple representation (see* (4.2i)*) of finite length, having a unique irreducible quotient representation $J(\delta \otimes \nu)$.*

2. *Every irreducible representation of G is of the form $J(\delta \otimes \nu)$; the triple (P, δ, ν) is unique up to conjugation by G.*

3. *The classes $\{[I(\delta \otimes \nu)]\}$ in $K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})$ form a basis; the transition matrix to the basis $\{[J(\delta \otimes \nu)]\}$ (of irreducible representations) is upper triangular.*

4. *For fixed δ, the continuous family of representations $\{I(\delta \otimes \nu) \mid \nu \in \widehat{A}\}$ all have the same image under* [gr]*:*

$$[\operatorname{gr} I(\delta \otimes \nu)] \simeq [\operatorname{gr} I(\delta \otimes \nu')] \qquad (\nu, \nu' \in \widehat{A}).$$

The change of basis matrix in (3) is explicitly computed by the Kazhdan-Lusztig theory of [17]. That is, we can compute explicit matrices of integers $M_{(\delta',\nu'),(\delta,\nu)}$ so that

$$[J(\delta \otimes \nu)] = \sum_{(\delta',\nu')} M_{(\delta',\nu'),(\delta,\nu)}[I(\delta', \nu')] \tag{7.14}$$

in the Grothendieck group $K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})$.

Computing the restriction to $K_{\mathbb{C}}$ of standard representations is relatively easy using the results of Knapp-Zuckerman in [8]. The conclusion is that for each (δ, ν) there is a (small) nonnegative integer $r = r(\delta, \nu)$ and a collection of 2^r irreducible representations $\tau_i \in \widehat{G}_{\mathrm{temp},\mathbb{R}}$ so that

$$I(\delta \otimes \nu)|_{K_{\mathbb{C}}} = \sum_{i=1}^{2^r} \tau_i. \tag{7.15}$$

Usually $r = 0$ and $\tau_1 = I(\delta \otimes 1)$.

Here is a summary of these facts, and some additional information about the equivariant K-theory of the nilpotent cone.

Theorem 7.16. *Suppose G is a real reductive group with maximal compact subgroup K. The homomorphism* [gr] *of* (7.12d) *is surjective, and factors to an isomorphism*

$$[\mathrm{gr}] \colon K(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})_{K_{\mathbb{C}}} \to K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*).$$

In particular, the classes $[\mathrm{gr}\,\tau]$, with τ a tempered irreducible representation of G having real infinitesimal character, are a basis of the equivariant K-theory $K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*)$. If $\pi \in \widehat{G}$, then the expression

$$[\mathrm{gr}\,\pi] = \sum_{\tau \in \widehat{G}_{temp,\mathbb{R}}} a_{\tau,\pi} [\mathrm{gr}\,\tau]$$

is explicitly computed by Kazhdan-Lusztig and Knapp-Zuckerman theory.

In a sense this theorem accomplishes the goal set forth at the end of Section 6: the invariant $AV(\pi)$ has been refined to the class $[\mathrm{gr}\,\pi]$ in equivariant K-theory, and this refined invariant is completely computed. But what we were seeking to understand (among other things) was the associated variety just as a set, and this we have not done at all. The reason is that the basis $[\mathrm{gr}\,\tau]$ of Theorem 7.16 is not nicely related to the support of coherent sheaves. In section 8 we will describe a second basis of $K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*)$ that *is* related to supports.

8 Equivariant K-theory and supports

In this section we will describe a basis for equivariant K-theory in the case that the algebraic group acts with finitely many orbits. It turns out that equivariant K-theory for (nearly) homogeneous spaces X is very closely related to the representation theory of algebraic groups. Here is why.

Suppose H is a complex algebraic group. Define

$$H^{\mathrm{red}} = H/(\text{unipotent radical of H}), \tag{8.1a}$$

the maximal reductive quotient of H. We write

$$\mathrm{Rep}(H) = \text{finite-dimensional algebraic representations of } H. \tag{8.1b}$$

The Grothendieck group of this category is called the *representation ring of* H, and written

$$R(H) =_{\text{def}} K\,\text{Rep}(H) \simeq K\,\text{Rep}(H^{\text{red}}) = R(H^{\text{red}}). \tag{8.1c}$$

If we write

$$\begin{aligned} \widehat{H} &= \text{equivalence classes of irr. alg. reps. of } H \\ &= \widehat{H^{\text{red}}}, \end{aligned} \tag{8.1d}$$

then elementary representation theory says that

$$R(H) = K\,\text{Rep}(H) \simeq \sum_{\rho \in \widehat{H}} \mathbb{Z}\rho. \tag{8.1e}$$

Tensor product of representations defines on the abelian group $R(H)$ the structure of a commutative ring, the *representation ring of* H.

Suppose X is a single point. Then a coherent sheaf on X is the same thing as a finite-dimensional vector space E; an H-equivariant structure is an algebraic representation ρ of H on E. That is, there is an equivalence of categories

$$\text{Coh}^H(\text{point}) \simeq \text{Rep}(H) \tag{8.1f}$$

Consequently

$$K^H(\text{point}) \simeq K\,\text{Rep}(H) = R(H) \simeq \sum_{\rho \in \widehat{H}} \mathbb{Z}\rho. \tag{8.1g}$$

More generally, suppose (ξ, V) is a finite-dimensional algebraic representation of H, and $P(V) = \mathcal{O}_V(V)$ is the algebra of polynomial functions on V. Then there is a natural inclusion

$$\text{Rep}(H) \hookrightarrow (P(V), H)\text{–Mod}^{\text{fg}} \simeq \text{Coh}^H(V), \quad E \mapsto E \otimes P(V). \tag{8.1h}$$

In contrast to (8.1f), this inclusion is very far from being an equivalence of categories (unless $V = \{0\}$). Nevertheless, just as in (8.1g), it induces an isomorphism in K-theory

$$K\,\text{Rep}(H) \simeq K^H(V); \tag{8.1i}$$

this is [16, Theorem 4.1]. (Roughly speaking, the image of the map (8.1h) consists of the *projective* $(P(V), H)$-modules; the isomorphism in K-theory arises from the existence of projective resolutions.)

Suppose now that $H \subset G$ is an inclusion of affine algebraic groups. Then there is an equivalence of categories

$$\text{Coh}^G(G/H) \simeq \text{Coh}^H(\text{point}) \simeq \text{Rep}(H); \tag{8.1j}$$

so the equivariant K-theory is

$$K^G(G/H) \simeq K^H(\text{point}) \simeq K\,\text{Rep}(H). \tag{8.1k}$$

More generally, if H acts on the variety Y, then we can form a fiber product

$$X = G \times_H Y$$

and calculate

$$\begin{aligned} \mathrm{Coh}^G(G \times_H Y) &\simeq \mathrm{Coh}^H(Y), \\ K^G(G \times_H Y) &\simeq K^H(Y). \end{aligned} \tag{8.1l}$$

If $Y \subset X$ is a closed H-invariant subvariety, then there is a right exact sequence

$$K^H(Y) \to K^H(X) \to K^H(X - Y) \to 0. \tag{8.1m}$$

This is established without the H for example in [4, Exercise 2.5.15], and the argument there carries over to the equivariant case.

Suppose now that the complex algebraic group H acts on the complex algebraic variety X with finitely many orbits:

$$X = \coprod_{i \in I} X_i = \coprod_{i \in I} H \cdot x_i \simeq \coprod_{i \in I} H/H_i, \tag{8.2a}$$

with I a finite set indexing the orbits. There is a natural partial order on I defined by the closure relation on orbits:

$$i \leq j \iff X_i \subset \overline{X_j}. \tag{8.2b}$$

Evidently this partial order is compatible with dimension:

$$i < j \implies \dim X_i < \dim X_j \iff \dim H_i > \dim H_j. \tag{8.2c}$$

Then

$$\overline{X_j} = \coprod_{i \leq j} X_i =_{\mathrm{def}} X_{\leq j} \tag{8.2d}$$

is a Zariski-closed H-stable subvariety of X. We will also be interested in the boundary

$$\partial \overline{X}_j = \overline{X_j} - X_j = \coprod_{i < j} X_i =_{\mathrm{def}} X_{<j}. \tag{8.2e}$$

The key result is

$$\text{if } X/H \text{ is finite, then } K_0^H(X) \text{ is a free } \mathbb{Z}\text{-module.} \tag{8.2f}$$

This is proven by induction on the number of orbits of H on X. If X is empty, then $K_0(X) = 0$ and there is nothing to prove. If X is not empty, we let

$$U = H/H' \subset X \tag{8.2g}$$

be an orbit of maximal dimension. Necessarily U is open, so $Y = X - U$ is a closed subvariety consisting of one less H orbit than X. The right exact sequence of (8.1m) extends to a long exact sequence in higher equivariant K-theory

$$\cdots \to K_1^H(H/H') \to K_0^H(Y) \to K_0^H(X) \to K_0^H(H/H') \to 0 \tag{8.2h}$$

([16, Theorem 2.7]). The proof of (8.1k) shows that $K_1^H(H/H')$ is the Quillen K_1 of a category of representations (of H'). This is (up to long exact sequences) a direct sum of categories of finite-dimensional complex vector spaces. By one of the fundamental facts about algebraic K-theory, it follows that $K_1^H(H/H')$ is a direct sum of copies of $\mathbb{C}^\times$, one for each irreducible representation of H'.

By inductive hypothesis, $K^H(Y)$ is free. Any group homomorphism from the divisible group $\mathbb{C}^\times$ to $\mathbb{Z}$ must be zero; so we get a short exact sequence

$$0 \to K^H(Y) \to K^H(X) \to K^H(H/H') \to 0 \tag{8.2i}$$

The first term is free by induction, and the last term is free with basis indexed by $\widehat{H'}$; so $K^H(X)$ is free, as we wished to show.

The argument proves that whenever H acts on X with finitely many orbits and $Y \subset X$ is closed and H-invariant, the maps in (8.1m) give a short exact sequence

$$0 \to K^H(Y) \hookrightarrow K^H(X) \twoheadrightarrow K^H(X - Y) \to 0 \tag{8.2j}$$

Suppose $(\xi, E_\xi) \in \widehat{H_j}$ is an irreducible algebraic representation, so that

$$\mathcal{E}_\xi = H \times_{H_j} E_\xi \tag{8.2k}$$

is an H-equivariant vector bundle on $X_j \simeq H/H_j$. We abuse notation by writing

$$\mathcal{E}_\xi = \text{sheaf of sections of } \mathcal{E}_\xi \in \operatorname{Coh}^H(X_j), \qquad [\mathcal{E}_\xi] \in K^H(X_j). \tag{8.2l}$$

We use the short exact sequence (8.2j)

$$0 \to K^H(\partial\overline{X_j}) \hookrightarrow K^H(\overline{X_j}) \twoheadrightarrow K^H(X_j) \to 0 \tag{8.2m}$$

Choose a preimage

$$[\widetilde{\mathcal{E}_\xi}] \in K^H(\overline{X}_j) \subset K^H(X), \qquad [\widetilde{\mathcal{E}_\xi}] \mapsto [\mathcal{E}_\xi] \in K^H(X_j); \tag{8.2n}$$

this is possible by (8.2m). In fact we can choose $\widetilde{\mathcal{E}_\xi}$ to be an H-equivariant coherent sheaf on $\overline{X}_j$; again we refer to the proof for coherent sheaves in [4, Exercise 2.5.15], which carries over to the equivariant case. Because of the exact sequence (8.2m), the class $[\widetilde{\mathcal{E}_\xi}]$ is unique up to a class in $K^H(\partial\overline{X}_j)$.

Proposition 8.3. *Suppose that the complex algebraic group H acts on the complex algebraic variety X with finitely many orbits; use notation as in* (8.2)*. Then the various classes*

$$[\widetilde{\mathcal{E}_\xi}] \in K^H(\overline{X}_j), \qquad X_j \simeq H/H_j, \quad \xi \in \widehat{H_j}$$

are a basis of $K^H(X)$. Each basis element $[\widetilde{\mathcal{E}_\xi}]$ is uniquely defined (by $\xi \in \widehat{H_j}$) modulo $K^H(\partial X_j)$.

Because of the geometric construction of this basis, and the left exactness of the sequences (8.2j), we deduce

Corollary 8.4. *Suppose that the complex algebraic group H acts on the complex algebraic variety X with finitely many orbits, and that $\mathcal{M}$ is a non-zero H-equivariant coherent sheaf on X. Write*

$$[\mathcal{M}] = \sum_{\xi} a_\xi [\widetilde{\mathcal{E}}_\xi], \tag{8.4a}$$

with ξ running over irreducible representations of the various orbit stabilizers H_j, and $a_\xi \in \mathbb{Z}$. Define

$$I(M) = \{\textit{maximal elements } j \in I \mid a_\xi \neq 0 \textit{ for some } \xi \in \widehat{H_j}\}. \tag{8.4b}$$

Then

1. *The associated variety*

$$V(\mathcal{M}) =_{def} \operatorname{Supp}(\mathcal{M}) = \bigcup_{j \in I(M)} \overline{X_j}. \tag{8.4c}$$

2. *For any $j \in I(M)$ and any $\xi \in \widehat{H_j}$, the coefficient a_ξ is a nonnegative integer independent of the choices of extensions $[\widetilde{\mathcal{E}}_{\xi'}]$.*

3. *Fix $j \in I(M)$ and $x_j \in X_j$ with isotropy group H_j. Then the genuine virtual representation $\chi(x_j, \mathcal{M})$ defined in* [18, Definition 2.12] *is*

$$\chi(x_j, \mathcal{M}) = \sum_{\xi \in \widehat{H_j}} a_\xi \xi. \tag{8.4d}$$

In particular, the multiplicity of the component $\overline{X_j}$ of $AV(\mathcal{M})$ is equal to

$$\sum_{\xi \in \widehat{H_j}} a_\xi \dim \xi, \tag{8.4e}$$

again by [18, Definition 2.12].

The corollary says that if we can compute the expansion of $[\mathcal{M}]$ in the geometric basis of Proposition 8.3, then we can compute the associated variety $AV(\mathcal{M})$, and even more, like the multiplicities of irreducible components. This is the idea we will use in Section 9 to compute associated varieties of Harish-Chandra modules.

According to the Weyl dimension formula (Theorem 3.6), the multiplicities in (8.4e) will be polynomial functions of the highest weights of the various ξ appearing.

9 Computing associated varieties

We return now to our real reductive group G with maximal compact subgroup K and complexification $K_{\mathbb{C}}$, in the setting of (7.12). We know (Theorem 6.5) that $K_{\mathbb{C}}$ acts on the nilpotent cone $\mathcal{N}_\theta^*$ with finitely many orbits $\{X_j \mid j \in I\}$, with isotropy subgroups $K_j \subset K_{\mathbb{C}}$. We have found two bases for the equivariant K-theory $K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*)$ of the nilpotent cone: a character-theoretic basis

$$\{[\mathrm{gr}\,\tau] \mid \tau \in \widehat{G}_{\mathrm{temp},\mathbb{R}}\} \tag{9.1a}$$

(Theorem 7.16), and a geometric basis

$$\{[\widetilde{\mathcal{E}}_\xi] \mid \xi \in \widehat{K_j}\} \tag{9.1b}$$

(Proposition 8.3). For any irreducible $(\mathfrak{g}_{\mathbb{C}}, K_{\mathbb{C}})$-module π, Kazhdan-Lusztig and Knapp-Zuckerman theory allow us to compute explicitly a formula

$$[\mathrm{gr}\,\pi] = \sum_{\tau \in \widehat{G}_{\mathrm{temp},\mathbb{R}}} M_{\tau',\pi}[\mathrm{gr}\,\tau'] \in K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*) \tag{9.1c}$$

(see Theorem 7.16). If on the other hand we can compute an expansion in the geometric basis

$$[\mathrm{gr}\,\pi] = \sum_{\xi \in \widehat{K_j}} a_{\xi,\pi}[\widetilde{\mathcal{E}}_\xi] \in K^{K_{\mathbb{C}}}(\mathcal{N}_\theta^*), \tag{9.1d}$$

then Corollary 8.4 computes the associated variety of π.

The final step in our journey will therefore be to find computable (change of basis) formulas

$$[\widetilde{\mathcal{E}}_\xi] = \sum_{\tau \in \widehat{G}_{\mathrm{temp},\mathbb{R}}} b_{\tau,\xi}[\mathrm{gr}\,\tau]. \tag{9.1e}$$

Once we have these formulas, we can invert them to get

$$[\mathrm{gr}\,\tau] = \sum_{\xi \in \widehat{K_j}} B_{\xi,\tau}[\widetilde{\mathcal{E}}_\xi]; \tag{9.1f}$$

and then plug these formulas into (9.1c) to get the explicit formulas (9.1d) that we need.

We will give only the barest sketch of how to get the change of basis formulas (9.1e). The ideas come from [1] and ultimately from [12]. It is convenient to work not with $\mathcal{N}_\theta^*$ but with the (equivariantly isomorphic) nilpotent cone

$$\mathcal{N}_\theta = \{e \in \mathfrak{s}_{\mathbb{C}} \mid e \in [\mathfrak{g}, \mathfrak{g}] \text{ nilpotent}\}. \tag{9.1g}$$

Here $\mathfrak{s}$ is defined in Proposition 2.2. According to [10] we can find elements h and f of $\mathfrak{g}_{\mathbb{C}}$ so that

$$h \in \mathfrak{k}_{\mathbb{C}}, f \in \mathfrak{s}_{\mathbb{C}}, \qquad [h,e] = 2e, \quad [h,f] = -2f, \quad [e,f] = h. \tag{9.1h}$$

The nonnegative eigenspaces of $\mathrm{ad}(h)$ define a θ-stable parabolic subalgebra

$$\mathfrak{q} = \mathfrak{l} + \mathfrak{u} = \mathfrak{g}_{\mathbb{C},0} + \mathfrak{g}_{\mathbb{C},\geq 1} \tag{9.1i}$$

with subscripts referring to $\mathrm{ad}(h)$ eigenvalues. The corresponding parabolic subgroup of $K_{\mathbb{C}}$ is

$$Q_{K_{\mathbb{C}}} = L_{K_{\mathbb{C}}} U_{K_{\mathbb{C}}} = K_{\mathbb{C}}^h \cdot \exp(\mathfrak{u} \cap \mathfrak{k}_{\mathbb{C}}; \tag{9.1j}$$

Proposition 9.2 (Kostant-Rallis [10])**.** *Suppose we are in the setting of* (9.1). *Write $U = \mathrm{Ad}(K_{\mathbb{C}})(e)$ for the corresponding nilpotent orbit.*

1. *The orbit $\mathrm{Ad}(Q_{K_{\mathbb{C}}})(e)$ is open and dense in $\mathfrak{s}_{\mathbb{C},\geq 2}$.*
2. *The centralizer of e in $K_{\mathbb{C}}$ is contained in $Q_{K_{\mathbb{C}}}$. Precisely,*
$$K_{\mathbb{C}}^e = L_{K_{\mathbb{C}}}^e U_{K_{\mathbb{C}}}^e,$$
and this is a Levi decomposition of the algebraic group.
3. *Define*
$$Z = K_{\mathbb{C}} \times_{Q_{K_{\mathbb{C}}}} \mathfrak{s}_{\mathbb{C},\geq 2},$$
a vector bundle over the projective variety $K_{\mathbb{C}}/Q_{K_{\mathbb{C}}}$. The natural map
$$\mu\colon Z \to \mathfrak{g}_{\mathbb{C}}, \quad \mu(Z) = \overline{U}, \qquad (k, S) \mapsto \mathrm{Ad}(k)S$$
is proper and birational.
4. *If $\mathcal{M}$ is a $K_{\mathbb{C}}$-equivariant coherent sheaf on Z, then each higher direct image sheaf $R^i\mu_*(\mathcal{M})$ is a $K_{\mathbb{C}}$-equivariant coherent sheaf on $\overline{U} \subset \mathcal{N}_\theta$. In this way we get a well-defined map*
$$\mu_*\colon K^{K_{\mathbb{C}}}(Z) \to K^{K_{\mathbb{C}}}(\overline{U}) \to K^{K_{\mathbb{C}}}(\mathcal{N}_\theta), \qquad [\mathcal{M}] \mapsto \sum_i (-1)^i [R^i\mu_*(\mathcal{M})].$$
5. *The isomorphisms* (8.1l) *and* (8.1i) *give*
$$K^{K_{\mathbb{C}}}(Z) \simeq K^{Q_{K_{\mathbb{C}}}}(\mathfrak{s}_{\mathbb{C},\geq 2}) \simeq R(Q_{K_{\mathbb{C}}}) \simeq R(L_{\mathbb{K}_{\mathbb{C}}}).$$
6. *For the nilpotent orbit $U = \mathrm{Ad}(K_{\mathbb{C}})(e)$,* (8.1k) *gives*
$$K^{K_{\mathbb{C}}}(U) \simeq R(Q_{K_{\mathbb{C}}}^e) \simeq R(L_{K_{\mathbb{C}}}^e).$$
7. *The surjective restriction map of* (8.1m) *is*
$$R(L_{K_{\mathbb{C}}}) \simeq K^{K_{\mathbb{C}}}(Z) \to K^{K_{\mathbb{C}}}(U) \simeq R(L_{K_{\mathbb{C}}}^e)$$
is just restriction of representations to $L_{K_{\mathbb{C}}}^e$.

8. *The map μ_* on K-theory defined in (4) restricts to the identity map on $K^{K_{\mathbb{C}}}(U)$, with U the common open orbit in Z and in $\overline{U}$.*

An amazing consequence of this proposition is that the restriction map

$$\mathrm{Res}\colon R(L_{K_{\mathbb{C}}}) \to R(L^e_{K_{\mathbb{C}}}) \simeq R(K^e_{\mathbb{C}}) \tag{9.3a}$$

is *surjective*: for every irreducible representation $\xi \in \widehat{K^e_{\mathbb{C}}}$, we can find a virtual representation $\sum n_i\sigma_i$ of $L_{K_{\mathbb{C}}}$ whose restriction to $L^e_{K_{\mathbb{C}}}$ is exactly ξ:

$$\xi = \sum n_i \sigma_i|_{L^e_{K_{\mathbb{C}}}} \qquad \sigma_i \in \widehat{L_{K_{\mathbb{C}}}}. \tag{9.3b}$$

This is a very unusual situation: for most algebraic subgroups of algebraic groups (as for most subgroups of finite groups) the restriction map on virtual representations is very far from surjective. The proof of this fact is not constructive; it involves finding a projective resolution. But in practice it seems not very difficult to find explicit formulas as in (9.3b).

If $(\sigma, F(\sigma)) \in \widehat{Q_{K_{\mathbb{C}}}}$, we get an equivariant vector bundle

$$\mathcal{F}(\sigma) \to K_{\mathbb{C}}/Q_{K_{\mathbb{C}}}, \tag{9.3c}$$

which we may pull back to an equivariant vector bundle (still called $\mathcal{F}(\sigma)$) on the bundle $Z \to K_{\mathbb{C}}/Q_{K_{\mathbb{C}}}$. The bundle Z and the map μ are closely related to the construction of Harish-Chandra modules by cohomological induction. The result is that there is an easily computable explicit formula

$$[\mu_*(\mathcal{F}(\sigma))] = \sum_{\tau' \in \widehat{G}_{\mathrm{temp},\mathbb{R}}} N_{\tau',\sigma}[\mathrm{gr}\,\tau'] \in K^{K_{\mathbb{C}}}(\mathcal{N}_\theta). \tag{9.3d}$$

(The `atlas` software, following [19], makes calculations exactly along these lines to compute restrictions to K of standard representations.

Given a formula (9.3b), we can define

$$[\widetilde{\mathcal{E}}_\xi] =_{\mathrm{def}} \sum_i n_i[\mu_*\mathcal{F}(\sigma_i)]. \tag{9.3e}$$

as our (virtual) coherent extension of the vector bundle $\mathcal{E}_\xi$ to the orbit closure. This leads to explicitly computable formulas

$$[\widetilde{\mathcal{E}}_\xi] = \sum_i \sum_{\tau' \in \widehat{G}_{\mathrm{temp},\mathbb{R}}} n_i N_{\tau',\sigma_i}[\mathrm{gr}\,\tau']. \tag{9.3f}$$

These are the formulas required in (9.1e) to complete our expression of $[\mathrm{gr}\,\pi]$ in the geometric basis, and therefore to compute $\mathrm{AV}(\pi)$.

References

[1] Pramod N. Achar, *On the equivariant K-theory of the nilpotent cone in the general linear group*, Represent. Theory **8** (2004), 180–211.

[2] Dan M. Barbasch and David A. Vogan Jr., *The local structure of characters*, J. Funct. Anal. **37** (1980), 27–55.

[3] N. Chriss and V. Ginzburg, *Representation Theory and Complex Geometry*, Birkhäuser, Boston-Basel-Berlin, 1997.

[4] Robin Hartshorne, *Algebraic geometry*, Springer, New York, 1977. Graduate Texts in Mathematics, No. 52.

[5] S. Helgason, *Groups and Geometric Analysis*, Mathematical Surveys and Monographs, vol. 83, American Mathematical Society, Providence, Rhode Island, 2000.

[6] R. Howe, *Wave front sets of representations of Lie groups*, Automorphic Forms, Representation Theory, and Arithmetic, 1981.

[7] Anthony W. Knapp, *Representation Theory of Semisimple Groups: An Overview Based on Examples*, Princeton University Press, Princeton, New Jersey, 1986.

[8] A. W. Knapp and Gregg J. Zuckerman, *Classification of irreducible tempered representations of semisimple groups*, Ann. of Math. (2) **116** (1982), no. 2–3, 389–501.

[9] ——, *Correction: "Classification of irreducible tempered representations of semisimple groups" [Ann. of Math. (2)* **116** *(1982), no. 2, 389–501*, Ann. of Math. (2) **119** (1984), no. 3, 639.

[10] Bertram Kostant and S. Rallis, *Orbits and representations associated with symmetric spaces*, Amer. J. Math. **93** (1971), 753–809.

[11] Robert P. Langlands, *On the classification of representations of real algebraic groups*, Representation Theory and Harmonic Analysis on Semisimple Lie Groups, 1989, pp. 101–170.

[12] W. McGovern, *Rings of regular functions on nilpotent orbits and their covers*, Invent. Math. **97** (1989), 209–217.

[13] George W. Mackey, *Unitary group representations in physics, probability, and number theory*, Mathematics Lecture Note Series, vol. 55, Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1978.

[14] Wilfried Schmid and Kari Vilonen, *Characteristic cycles and wave front cycles of representations of reductive Lie groups*, Ann. of Math. (2) **151** (2000), no. 3, 1071–1118.

[15] Jirō Sekiguchi, *Remarks on real nilpotent orbits of a symmetric pair*, J. Math. Soc. Japan **39** (1987), no. 1, 127–138.

[16] R. W. Thomason, *Algebraic K-theory of group scheme actions*, Algebraic topology and algebraic K-theory (Princeton, N.J., 1983), Ann. of Math. Stud., vol. 113, Princeton Univ. Press, Princeton, NJ, 1987, pp. 539–563.

[17] David A. Vogan Jr., *Irreducible characters of semisimple Lie groups III. Proof of the Kazhdan-Lusztig conjectures in the integral case*, Invent. Math. **71** (1983), 381–417.

[18] ______, *Associated varieties and unipotent representations*, Harmonic Analysis on Reductive Groups, 1991, pp. 315–388.

[19] ______, *Branching to a maximal compact subgroup*, Harmonic Analysis, Group Representations, Automorphic Forms, and Invariant Theory: In Honour of Roger E. Howe, 2007.