2, 3, 4, or 6. Comparison of lists (2.10.3) and (2.10.4) reveals it is exactly the Coxeter graphs which have an m_{rs} other than 2, 3, 4, or 6 which do not survive to become Dynkin diagrams.

7. This is for "reduced" root systems, which is what is encountered in classifying simple complex Lie algebras. For real Lie algebras, nonreduced root systems, e.g., BC_n, can also occur [**Helg2, Serr1**].

3. Representation theory. Research into representations (actions on vector spaces via linear transformations) of Lie groups, motivated on one hand by physics [**FlSz, Mack1, ITGT1–17, Barg3**] and on the other by the theory of automorphic forms [**GGPS, JaLa, Weil1, BoCa**] with deep roots in classical analysis and with strong ties to differential equations, and of course also propelled by its internal dynamics, has been a major part of the mathematical enterprise since roughly World War II. Considering the diversity of motivations, goals, people, and methods involved, the subject displays a remarkable amount of unity. A major source of the unity is the philosophy of the *orbit method* (also known by the more fashionable term *geometric quantization* [**Blat, Kiri, Kost1, Sour**]). Although we can only sample from the wide range of results that have been established, the overall coherence provided by the viewpoint of the orbit method allows us to convey much more of the subject than would otherwise be possible. An interesting technical point, however, is that the orbit method is almost exclusively a method of *interpretation*, a way of organizing results into a coherent (and often very beautiful) pattern. It provides little in the way of technical tools for proofs or computations. Thus, for example, several of the major results of Harish-Chandra on representations of semisimple groups have found elegant interpretations in terms of the orbit method [**Ross1, 2, DuVe, DuHV**]. However, these interpretations have provided no short-cuts to Harish-Chandra's proofs of these results.

A proper discussion of representation theory requires an aggravatingly long technical preparation. We are going to try to ignore that here. For the convenience of the reader, basic definitions and constructions have been summarized in Appendix 1. The discussion below refers to Appendix 1 as necessary. The reader who finds these references too distracting may wish to acquaint himself, at least in a rough way, with Appendix 1 before reading the main body of this section.

3.1. *An example: the quantum harmonic oscillator.* To illustrate the potential uses of representation theory, and its attraction, I can produce no better example than the spectral analysis of the quantum mechanical harmonic oscillator. This is elementary almost to the point of simple-mindedness, yet it contains the seeds of extremely varied developments that form subjects of active current research. In particular, it is basic for the orbit method to be discussed later. Also, it exhibits the extreme elegance of the best Lie algebraic computations.

3.1.1. A quantum mechanical system is defined by a selfadjoint operator called the *Hamiltonian operator* on a Hilbert space $\mathscr{H}$ **[Mack3]**. Analysis of the system involves describing the spectral decomposition, especially the eigenvalues and eigenvectors, of the Hamiltonian. For the one-dimensional quantum harmonic oscillator, the Hilbert space is $L^2(\mathbf{R})$, and the Hamiltonian is **[Shan]**

$$T = \frac{d^2}{dx^2} - x^2. \tag{3.1.1.1}$$

To find the spectrum of T, consider the operators p, q on $L^2(\mathbf{R})$ defined by

$$p(f)(x) = \frac{df}{dx}(x), \qquad q(f)(x) = ixf(x) \tag{3.1.1.2}$$

for f sufficiently nice in $L^2(X)$. It is easy to check that the four operators T, p, q, and 1, the identity operator, span a four-dimensional Lie algebra: the commutators

$$[A, B] = AB - BA \tag{3.1.1.3}$$

of two of these operators is a linear combination of some or all of them. Indeed, easy computations show

$$\begin{aligned} &\text{(a) } [p, q] = i, \\ &\text{(b) } [T, p] = -2iq, \ [T, q] = 2ip, \end{aligned} \tag{3.1.1.4}$$

and of course the commutator of 1 with anything is zero.

Let us set

$$\mathbf{a} = \frac{d}{dx} + x = p - iq, \qquad \mathbf{a}^+ = \frac{d}{dx} - x = p + iq. \tag{3.1.1.5a}$$

Then we observe

$$[\mathbf{a}^+, \mathbf{a}] = 2, \tag{3.1.1.5b}$$

$$\mathbf{a}^+ = -\mathbf{a}^*, \tag{3.1.1.5c}$$

where $\mathbf{a}^*$ indicates the operator on $L^2(\mathbf{R})$ adjoint to $\mathbf{a}$, and

$$T = \tfrac{1}{2}(\mathbf{a}^+\mathbf{a} + \mathbf{a}\mathbf{a}^+). \tag{3.1.1.5d}$$

Further we can see that the vector

$$v_0 = e^{-x^2/2} \tag{3.1.1.5e}$$

is annihilated by $\mathbf{a}$:

$$\mathbf{a}v_0 = 0. \tag{3.1.1.5f}$$

Now let us forget we are dealing with specific operators on $L^2(\mathbf{R})$. Let us simply suppose we have some Hilbert space on which are defined two operators, $\mathbf{a}$ and $\mathbf{a}^+$, satisfying relations (3.1.1.5b,c), such that there is a vector v_0 annihilated by the operator $\mathbf{a}$. Define

$$v_j = (\mathbf{a}^+)^j(v_0) = \mathbf{a}^+(v_{j-1}), \qquad j = 1, 2, 3, \ldots. \tag{3.1.1.6}$$

I claim

$$\mathbf{a}(v_j) = -2jv_{j-1}. \tag{3.1.1.7}$$

This may be easily verified by use of the commutator identity

$$[\mathbf{a}, (\mathbf{a}^+)^k] = \sum_{j=0}^{k-1}(\mathbf{a}^+)^j[\mathbf{a}, \mathbf{a}^+](\mathbf{a}^+)^{k-j-1} = -2k(\mathbf{a}^+)^{k-1}.$$

Using (3.1.1.6) and (3.1.1.7) we can verify that, if T is defined by formula (3.1.1.5d) then

$$T(v_j) = -(2j+1)v_j. \tag{3.1.1.8}$$

Thus the v_j are eigenvectors for T. Since T is selfadjoint, this means the v_j are mutually orthogonal. We can even determine the Hilbert space norms of the v_j's. If the inner product is denoted by (,) we can compute

$$(v_j, v_j) = (\mathbf{a}^+v_{j-1}, \mathbf{a}^+v_{j-1}) = -(\mathbf{a}\mathbf{a}^+v_{j-1}, v_{j-1}) = 2j(v_{j-1}, v_{j-1}).$$

Hence

$$(v_j, v_j) = 2^j\, j!(v_0, v_0). \tag{3.1.1.9}$$

It follows that if we put

$$u_j = (2^j\, j!(v_0, v_0))^{-1/2}v_j, \tag{3.1.1.10}$$

then the u_j form an orthogonal sequence of eigenvectors for T, and

$$\mathbf{a}u_j = -(2j)^{1/2}u_{j-1}, \qquad \mathbf{a}^+u_j = (2(j+1))^{1/2}u_{j+1}. \tag{3.1.1.11}$$

If we now return to the concrete situation which gave rise to equations (3.1.1.5), we see that the commutation relations (3.1.1.4) (which follow from (3.1.1.5a,b,d) allow us to construct what can be shown to be an orthonormal eigenbasis for T, and in particular to determine its spectrum.

3.1.2. The structure revealed by the calculations above has significance far beyond its application to the determination of the spectrum of the harmonic oscillator. In particular, the commutation relations (3.1.1.4a) between p and q, or (3.1.1.5a) between $\mathbf{a}$ and $\mathbf{a}^+$, which are known as Heisenberg's *Canonical Commutation Relations* (CCR for short) (cf. **[Mack3, Shan, Weyl3]**, etc.), have been found to be fundamental to quantum mechanics. They imply the uncertainty principle, which asserts that no particle state (i.e., vector in $L^2(\mathbf{R})$) can exist for which momentum and position are simultaneously well defined (i.e., which is a simultaneous eigenvector for p, the "momentum operator," and q, the "position operator"). See **[DyMc, Foll1, Körn, Shan]**, etc.

Further, equation (3.1.1.7) shows that a triple $(\mathbf{a}, \mathbf{a}^+, v_0)$ consisting of two operators $\mathbf{a}, \mathbf{a}^+$ satisfying (3.1.1.5b,c) together with a vector v_0 satisfying (3.1.1.5f) is essentially unique. This may be taken as a version of another foundational result of quantum mechanics, the Stone-von Neumann Theorem (cf. Theorem 3.3.2.4 and **[Cart, Foll, Mack3, Howe4, vNeu]**, etc.), which asserts the uniqueness, under appropriate technical hypotheses, of the canonical commutation relations. (We note that some sort of condition, such as (3.1.1.5f), is needed to supplement the CCR (3.1.1.5a) in order to guarantee uniqueness. The possibilities for nonuniqueness were exploited by J. Bernstein to obtain interesting results in distribution theory **[Bern1, Bern2, Borl2]**.)

3.1.3. The uniqueness result of §3.1 has an easy extension to larger systems of operators. Let $\{p_j, q_j\}_{j=1}^n$ be a collection of $2n$ operators satisfying the following relations (known again as the Canonical Commutation Relations):

$$[p_j, p_k] = 0 = [q_j, q_k] \qquad [p_j, q_k] = i\,\delta_{jk}. \tag{3.1.3.1}$$

Then the p's and q's, together with 1, the identity operator, span a $(2n+1)$-dimensional Lie algebra, now widely known as the *Heisenberg Lie algebra*. The Heisenberg algebra may be realized on $L^2(\mathbf{R}^n)$ by taking q_j to be multiplication by ix_j and p_j to be partial differentiation with respect to x_j. The Stone von-Neumann Theorem applies also to these systems and asserts, again under some natural hypotheses, that the realization of the p's and q's by ix_j's and $\frac{\partial}{\partial x_j}$'s is essentially unique. One form of this result amounts to a classification of the irreducible unitary representations (see §A.1.7) of a certain nilpotent Lie group, known as the Heisenberg group (see §3.3 and also **[Cart, Foll, Howe4, Moor]**, etc.). This is a basic step in the classification of the unitary dual (see §A.1.7) of nilpotent and solvable Lie groups **[AuKo, Kiri, Moor, Puka3]**.

The Heisenberg Lie algebra is closely connected not only with the harmonic oscillator, but with many other important equations of physics, both classical and quantum **[Sthr, Howe6, Engl]**. Extended to infinite numbers of variables, it plays a key role in quantum field theory **[Segal, Shal, Thir]** and the theory of "loop groups" and vertex algebras **[Garl, FrLm, FrKa, Kac1–7, KaPe, Lepo1, Lepo2]**.

In addition to these applications to physics, mathematical structures attached to the CCR are important in algebraic geometry (invariant theory **[Howe1]**, abelian varieties **[Cart, Igus, Mumf]**), number theory (theory of θ-series **[Cart, Gelb2, Howe5, HoPS, KuMi1, 2, 3, LiVe, ToWa1, 2]**, etc., K-theory **[Rama]**), and differential equations (Hamiltonian systems **[Olve]** (cf. §3.2), pseudo-differential and Fourier integral operators **[FePh, Foll1, GuSt1, Howe3, 4]**, several complex variables **[Foll2, FoSt, Stan]**, and D-modules

[**Borl1**, **Bern1**, **Bern2**]). Some of these topics will be touched on in the discussion which follows.

3.2. *The orbit method.* The philosophy which describes a large portion of the representation theory of Lie groups is a descendant of the correspondence principle of early quantum mechanics [**Bohr**, **Iken**, **Jamm**]. Since it is a philosophy and not a theorem, it is difficult to formulate in such a way that is not clearly false in some cases, but still appears to have content. But roughly the idea is that, if G is a connected Lie group, then for each "classical dynamical system" for G, there should be a corresponding "quantum dynamical system," which would be a unitary representation.

3.2.1. What could this mean? The key to the matter is symplectic geometry [**AbMa**, **Grom**, **GuSt**, **Wein**]. This is geometry based on a skew-symmetric bilinear form, in contrast to Euclidean or Riemannian geometry, which is based on a symmetric bilinear form. It is a slippery, less tangible kind of geometry; there is no notion of "distance" or "angles" in symplectic geometry. However, somewhat latterly because of its elusive nature, symplectic geometry has come to be seen to be of fundamental importance. Lie theory in particular seems to be steeped in symplecticism, owing to the anti-symmetry of the Lie bracket.

Let V be a finite-dimensional real vector space. A *symplectic form* $\langle\ ,\ \rangle$ on V is a nondegenerate skew-symmetric bilinear form. Nondegeneracy means that the map $\alpha\colon V \to V^*$ defined by

$$\alpha(v)(v') = \langle v', v\rangle, \qquad v, v' \in V, \tag{3.2.1.1}$$

is an isomorphism. Standard elementary arguments [**Lang3**, **Jaco2**] show that for V to have a symplectic form, V must have even dimension, say $2n$. Further, given n, there is essentially just one symplectic form. Precisely, we can, again by very elementary arguments, always find a *symplectic basis* for V, that is, a basis $\{e_i, f_i\}_{1\leq i\leq n}$, such that

$$\langle e_i, e_j\rangle = 0 = \langle f_i, f_j\rangle, \quad \langle f_i, e_j\rangle = \delta_{ij}, \qquad 1 \leq i, j \leq n, \tag{3.2.1.2}$$

where δ_{ij} is Kronecker's delta. If x_i, y_i are the coordinates with respect to the symplectic basis (we call them *symplectic coordinates*), then

$$\langle v, v'\rangle = \sum_{i=1}^{n} x_i' y_i - x_i y_i'. \tag{3.2.1.3}$$

From a symplectic form on V, we can construct a Lie algebra structure on $C^\infty(V, \mathbf{R})$, the real-valued smooth functions on V; the Lie bracket in this case is known as the *Poisson bracket*. In formulas, in the coordinates of (3.2.1.3), we have

$$\{P, Q\} = \sum_{i=1}^{n} \frac{\partial P}{\partial y_i}\frac{\partial Q}{\partial x_i} - \frac{\partial P}{\partial x_i}\frac{\partial Q}{\partial y_i}, \qquad P, Q \in C^\infty(V). \tag{3.2.1.4}$$

There are at least three conceptual ways of thinking about this formula. Much of the richness of Hamiltonian mechanics stems from the fact that they all yield the same answer, formula (3.2.1.4).

First, recall that the derivative or differential

$$dP = \sum_{i=1}^{n} \frac{\partial P}{\partial x_i}\, dx_i + \frac{\partial P}{\partial y_i}\, dy_i \tag{3.2.1.5}$$

is a function on V with values in V^*. We know the symplectic form defines an isomorphism α from V to V^*. Thus we can consider $\alpha^{-1}(dP)$ and $\alpha^{-1}(dQ)$, which are V-valued functions on V. We can compute

$$\alpha^{-1}(dP) = \sum_{i=1}^{n} \frac{\partial P}{\partial y_i}\, e_i - \frac{\partial P}{\partial x_i}\, f_i \tag{3.2.1.6}$$

and from this that the Poisson bracket may be expressed as

$$\{P, Q\} = \langle \alpha^{-1}(dP), \alpha^{-1}(dQ)\rangle. \tag{3.2.1.7}$$

Second, we can regard the V-valued function $\alpha^{-1}(dP)$ as defining a vector field on V. (Indeed, this is the correct thing to do from the point of view of differential geometry.) We can then differentiate a function with respect to $\alpha^{-1}(dP)$. The Poisson bracket can also be expressed in these terms:

$$\{P, Q\} = \alpha^{-1}(dP)(Q). \tag{3.2.1.8}$$

Third, if we think of both $\alpha^{-1}(dP)$ and $\alpha^{-1}(dQ)$ as vector fields, then we can consider their Lie bracket, as in formula (2.3.8), and we have

$$[\alpha^{-1}(dP), \alpha^{-1}(dQ)] = \alpha^{-1}(d\{P, Q\}). \tag{3.2.1.9}$$

This formula shows that the map

$$P \to \alpha^{-1}(dP) \tag{3.2.1.10}$$

is a Lie algebra homomorphism from $C^\infty(V; \mathbf{R})$, equipped with the Poisson bracket, to the space of vector fields on V, with their natural Lie bracket.

This third interpretation of Poisson bracket leads one to ask what the image of the map (3.2.1.10) looks like. From the form (3.2.1.6) of $\alpha^{-1}(dP)$ it is clear that it cannot be an arbitrary vector field; its coefficients must satisfy the obvious "integrability conditions" imposed by the equality of mixed partial derivatives, namely, if we write a vector field

$$v = \sum a_i e_i + b_i f_i, \tag{3.2.1.11a}$$

then if $v = \alpha^{-1}(dP)$ for some $P \in C^\infty(V; \mathbf{R})$ we must have

$$\frac{\partial a_i}{\partial y_j} = \frac{\partial a_j}{\partial y_i}, \quad \frac{\partial a_i}{\partial x_j} = -\frac{\partial b_j}{\partial y_i}, \quad \frac{\partial b_i}{\partial x_j} = \frac{\partial b_j}{\partial x_i}. \tag{3.2.1.11b}$$

Conversely, the Poincaré Lemma **[Gold, Ster]** tells us the conditions (3.2.1.11b) do guarantee that v will be of the form $\alpha^{-1}(dP)$. But of more interest are the following equivalent geometric interpretations of the integrability conditions.

PROPOSITION 3.2.1.12. *A vector field v is in the image of map* (3.2.1.10) *if and only if*

(i) *the natural action of v on $C^\infty(V)$ is a derivation of the Poisson bracket, i.e.,*

$$v(\{P, Q\}) = \{v(P), Q\} + \{P, v(Q)\}, \tag{3.2.1.13}$$

or

(ii) *the natural action of v on exterior forms on V annihilates the form*

$$\omega = \sum_{i=1}^{n} dx_i \wedge dy_i. \tag{3.2.1.14}$$

In terms of the one-parameter group φ_t (or local group) generated by v, as described in §2.1, condition (3.2.1.13) says that the φ_t will be automorphisms of the Poisson bracket, and the equivalent condition (3.2.1.14) says the φ_t will preserve the differential form ω. Clearly the diffeomorphisms satisfying either of these conditions will form a group, which is sometimes called the group of *symplectomorphisms.* (A more traditional term is *canonical transformation.*) Roughly speaking, the vector fields satisfying the equivalent conditions of Proposition 3.2.1.12 form the Lie algebra of this group; consequently we will denote the space of them by $\mathrm{Vect}_{\mathrm{Sp}}(V)$. This allows us to summarize the discussion just above by saying the map (3.2.1.10) takes $C^\infty(V; \mathbf{R})$ to $\mathrm{Vect}_{\mathrm{Sp}}(V)$.

An important technical point about the map (3.2.1.10) is that it is almost but not quite an isomorphism: it has a one-dimensional kernel, consisting of the constant functions. Also, it is easy to check from formula (3.2.1.4) (by letting P be a fixed function, and letting Q vary through the coordinate functions x_i, y_i) that the constants are precisely the center of the Lie algebra $C^\infty(V; \mathbf{R})$ with Poisson bracket. Thus we have an exact sequence

$$0 \to \mathbf{R} \to C^\infty(V; \mathbf{R}) \xrightarrow{\alpha^{-1}\circ d} \mathrm{Vect}_{\mathrm{Sp}}(V) \to 0 \tag{3.2.1.15}$$

which exhibits $C^\infty(V; \mathbf{R})$ as a one-dimensional central extension of $\mathrm{Vect}_{\mathrm{Sp}}(V)$.

To illustrate the difference the central extension (3.2.1.15) makes, consider the Lie algebra generated by the coordinate functions x_i, y_i. It is easy to check that

$$\alpha^{-1}(dx_i) = -f_i, \qquad \alpha^{-1}(dy_i) = e_i. \tag{3.2.1.16}$$

Hence the vector fields $\alpha^{-1}(d\lambda)$, $\lambda \in V^*$, are just the directional derivatives on V; they form an abelian Lie algebra, whose corresponding group is just V, acting on itself by translations. However, under Poisson bracket, the x_i and y_i generate a nonabelian Lie algebra: we have

$$\{x_i, x_j\} = 0 = \{y_i, y_j\}, \qquad \{y_j, x_i\} = \delta_{ij}. \tag{3.2.1.17}$$

These are simply a version of the CCR (see (3.1.3.1); the normalization here is slightly different from (3.1.3.1)).

Hence the Lie algebra generated by V^* under $\{,\}$ is a $(2n+1)$-dimensional, two-step nilpotent Heisenberg algebra

$$\mathbf{h}(V) = V^* \oplus \mathbf{R}. \tag{3.2.1.18}$$

Although when realized via the Poisson bracket, the Heisenberg Lie algebra is described in terms of V^*, it is more natural to describe it in terms of V, which is easy to do since we have identified V and V^* via the map α of formula (3.2.1.1). Thus we prefer to write

$$\mathbf{h}(V) = V \oplus \mathbf{R}. \tag{3.2.1.19a}$$

Then the Lie bracket looks like

$$[(v, t), (v', t')] = (0, \langle v, v'\rangle), \qquad v, v' \in V, t, t' \in \mathbf{R}. \tag{3.2.1.19b}$$

Finally, to conclude this subsection, we note that the space $S^2(V^*)$ of homogeneous quadratic polynomials forms a Lie algebra under the Poisson bracket. This algebra normalizes the Heisenberg Lie algebra $\mathbf{h}(V)$ discussed just above, and via the map (3.2.1.10) it is sent isomorphically to the Lie algebra $\mathbf{sp}(V)$ of the symplectic group $\mathrm{Sp}(V)$ of linear transformations of V which preserve $\langle\ ,\ \rangle$. (See §3.5.5 for more discussion of this remarkable realization of $\mathbf{sp}(V)$.)

3.2.2. We can use the discussion of §3.2.1 to define a *symplectic manifold* M in a manner entirely analogous to the usual definition ([**Gold**, **Helg2**, **AbMa**, **Ster**], etc.) of smooth manifold: one covers the underlying point set of the manifold M with local coordinate patches, such that the local coordinate functions are the coordinates with respect to a standard symplectic basis of a symplectic vector space; instead of letting the coordinate changes on overlapping charts be arbitrary diffeomorphisms, one requires them to be symplectomorphisms. Then if one interprets Proposition 3.2.1.12 using the standard language of differentiable manifolds (see references just above), one sees M has the following properties:

(3.2.2.1)

(i) There is a distinguished closed exterior 2-form ω on M, i.e., a section of $\Lambda^2 T^*(M)$, with the property that the alternating bilinear form induced by ω on the tangent space at each point of M is a symplectic form. (The 2-form ω will have the form (3.2.1.14) in each local chart.)

(ii) The space $C^\infty(M)$ is endowed with a Lie algebra structure, called the Poisson bracket, and denoted $\{\ ,\ \}$. This will satisfy the appropriately coordinate-free versions of properties (3.2.1.7), (3.2.1.8), and (3.2.1.9). (On each coordinate patch, the bracket $\{,\}$ will be given by formula (3.2.1.4).)

Alternately, one could define a symplectic manifold M as one having a distinguished closed 2-form, as in (3.2.2.1)(i), or as having a Poisson bracket structure on $C^\infty(M; \mathbf{R})$, as in (3.2.2.1)(ii). Some basic lemmas (Darboux's Theorem) then guarantee that M can be covered by local coordinate charts, in the way we imagined to begin with ([**AbMa, Olve, Ster**]).

In any case, the Poisson bracket gives us a homomorphism of Lie algebras

$$C^\infty(M; \mathbf{R}) \to \mathrm{Vect}_{\mathrm{Sp}}(M),$$

where again $\mathrm{Vect}_{\mathrm{Sp}}(M)$ is the Lie algebra of vector fields which generate (local) one-parameter groups of symplectomorphisms. The kernel of the map is the space of locally constant functions on M. Since the characteristic functions of the connected components of M form a canonical basis for this space, we may identify it with the 0th cohomology group $H^0(M)$. Also, we have seen that via the map α of formula (3.2.1.1), the space $\mathrm{Vect}_{\mathrm{Sp}}(M)$ is identified with the closed 1-forms on M, and the map from $C^\infty(M; \mathbf{R})$ is simply exterior differentiation. Hence the cokernel of this map is identified to the first deRham cohomology group $H^1(M)$. Thus we have an exact sequence

$$(3.2.2.2) \qquad 0 \to H^0(M) \to C^\infty(M; \mathbf{R}) \to \mathrm{Vect}_{\mathrm{Sp}}(M) \to H^1(M) \to 0.$$

There are three main sources of examples of symplectic manifolds.

(a) *Cotangent bundles*: If M is any manifold, then $T^*(M)$, the cotangent bundle of M, is in a natural way a symplectic manifold [**AbMa**, **Blat**, **Ster**].

(b) *Kähler manifolds* [**LaBe**, **Hart**, **Weil3**]: Let U be a complex vector space, and let $(\ ,\ )$ be a Hermitian inner product on U. Then the imaginary part of $(\ ,\ )$ defines a symplectic form on the real vector space obtained from U by restricting scalars. A Kähler manifold is a complex manifold M which is endowed with a Hermitian metric on its holomorphic tangent bundle, whose imaginary part is a closed $(1, 1)$-form, and which thus defines a symplectic structure on M. Kähler manifolds are significant because they include all nonsingular projective algebraic varieties: complex projective space $\mathbf{CP}^n$ possesses a Kähler metric, the Fubini-Study metric [**GrHa**], the unique metric invariant under the action of the unitary group U_{n+1} on $\mathbf{CP}^n$; and any nonsingular projective subvariety of $\mathbf{CP}^n$ inherits this metric by restriction. For purposes of obtaining symplectic manifolds, one can equally well consider "pseudo-Kähler" manifolds, defined in the same way as Kähler manifolds, except the Hermitian "metric" need not be positive definite.

(c) *Coadjoint orbits*: For us, this is the most important class of examples. Let G be a Lie group, write $\mathrm{Lie}(G) = \mathbf{g}$, and let $\mathbf{g}^*$ be the dual space to $\mathbf{g}$. The group G acts on $\mathbf{g}$ via Ad, the adjoint action, and therefore acts on $\mathbf{g}^*$ via the contragredient to Ad, called the *coadjoint action*, and denoted Ad^*. Consider $\lambda \in \mathbf{g}^*$. Let

$$(3.2.2.3) \qquad R_\lambda = \{g \in G : \mathrm{Ad}^* g(\lambda) = \lambda\}$$

be the stabilizer or isotropy group of λ, the subgroup of G which leaves λ fixed. Its Lie algebra is

$$\mathbf{r}_\lambda = \{x \in \mathbf{g} : \mathrm{ad}^*(x)(\lambda) = 0\}. \tag{3.2.2.4}$$

The map

$$e_\lambda : g \to \mathrm{Ad}^* g(\lambda) \tag{3.2.2.5}$$

defines a surjective, G-equivariant map from the coset space G/R_λ to

$$\mathscr{O}_\lambda = \{\mathrm{Ad}^* g(\lambda) : g \in G\}, \tag{3.2.2.6}$$

the $\mathrm{Ad}^* G$ orbit through λ. Differentiating the map e_λ at the origin gives an isomorphism

$$\mathbf{g}/\mathbf{r}_\lambda \simeq T(\mathscr{O}_\lambda)_\lambda \tag{3.2.2.7}$$

of the quotient $\mathbf{g}/\mathbf{r}_\lambda$ with the tangent space to $\mathscr{O}_\lambda$ at λ.

Consider on $\mathbf{g}$ the antisymmetric bilinear form

$$\langle x , y \rangle_\lambda = \lambda([x , y]). \tag{3.2.2.8}$$

One can easily check that the radical of the form $\langle \ , \ \rangle_\lambda$—defined as

$$\{x \in \mathbf{g} : \langle x , y \rangle_\lambda = 0 \text{ for all } g \in \mathbf{g}\},$$

that is, the vectors which are orthogonal to everything with respect to the form $\langle \ , \ \rangle_\lambda$ on $\mathbf{g}$—is precisely $\mathbf{r}_\lambda$. Hence the form $\langle \ , \ \rangle_\lambda$ factors to define a non-degenerate form on the quotient $\mathbf{g}/\mathbf{r}_\lambda$. In view of the isomorphism (3.2.2.7), we can push $\langle \ , \ \rangle_\lambda$ forward to define a symplectic form on the tangent space $T(\mathscr{O}_\lambda)_\lambda$ to $\mathscr{O}_\lambda$ at λ. Since this can be done at every point of $\mathbf{g}^*$, and since it is a canonical construction, this will produce a G-invariant differential 2-form which induces a symplectic form on the tangent space to $\mathscr{O}_\lambda$ at every point. A computation shows **[GuSt, AbMa]** that this canonically defined 2-form is in fact closed. (It should not be surprising that this is essentially a consequence of the Jacobi identity.) Hence $\mathscr{O}_\lambda$ is a symplectic manifold; further G acts transitively on $\mathscr{O}_\lambda$ via symplectomorphisms.

Some coadjoint orbits are isomorphic to cotangent bundles, and others support Kähler or pseudo-Kähler metrics.

Lie [**LiEn**, vol. 2, p. 294] was apparently aware of the symplectic structure on coadjoint orbits, or at least the associated Poisson bracket, but it was subsequently forgotten until the 1960s when its importance for representation theory was appreciated [**Bere**, **Blat**, **Kiri**, **Kost1**].

3.2.3. Let G be a Lie group and let M be a connected symplectic manifold. Suppose G acts on M by symplectomorphisms. Differentiating the action of G yields a homomorphism β from $\mathrm{Lie}(G)$ to $\mathrm{Vect}_{\mathrm{Sp}}(M)$. Denote the image of $\mathrm{Lie}(G)$ in $\mathrm{Vect}_{\mathrm{Sp}}(M)$ by $\mathbf{g}$. We would like to lift $\mathbf{g}$ to a subalgebra of $C^\infty(M)$. According to the sequence (3.2.2.2) there are two obstructions to doing this. The first is that $\mathbf{g}$ may not be in the image of the map from $C^\infty(M)$ to $\mathrm{Vect}_{\mathrm{Sp}}(M)$, that is, some elements of $\mathbf{g}$ may represent nontrivial cohomology in $H^1(M)$. If M is simply connected, then $H^1(M) = 0$ **[Mass]**, so we can eliminate this obstruction by passing to a covering of M if necessary. So suppose $\mathbf{g}$ is in the image of $C^\infty(M)$. Denote the inverse image of $\mathbf{g}$ in $C^\infty(M)$, via the sequence (3.2.2.2), by $\tilde{\mathbf{g}}$. Then we have a diagram:

$$\begin{tikzcd} & & & \mathrm{Lie}(G) \arrow[d, "\beta"'] & \\ 0 \arrow[r] & \mathbf{R} \arrow[r] & \tilde{\mathbf{g}} \arrow[r] & \mathbf{g} \arrow[r] & 0 \end{tikzcd}$$

The Lie algebra $\tilde{\mathbf{g}}$ is a central extension of $\mathbf{g}$ by $\mathbf{R}$, and thus defines a certain cohomology class γ in $H^2(\mathbf{g}; \mathbf{R})$ (see **[Jaco1, Kost1]**). We can lift the homomorphism β to a homomorphism

$$\hat{\beta}\colon \mathrm{Lie}(G) \to \tilde{\mathbf{g}} \subseteq C^\infty(M; \mathbf{R})$$

if and only if the pullback $\beta^*(\gamma) \in H^2(\mathrm{Lie}(G); \mathbf{R})$ vanishes. If this happens, then there is a choice of liftings $\tilde{\beta}$ of β, corresponding to the homomorphisms of Lie(G) to $\mathbf{R}$ (which form the group $H^1(\mathrm{Lie}(G); \mathbf{R}) \simeq (\mathbf{g}/\mathbf{g}^{(2)})^*$).

By a *Hamiltonian action* β of G on M, we mean an action of G on M by symplectomorphisms, together with a compatible homomorphism

$$\tilde{\beta}\colon \mathrm{Lie}(G) \to C^\infty(M; \mathbf{R})$$

such that the diagram

(3.2.3.1)
$$\begin{tikzcd} & & & \mathrm{Lie}(G) \arrow[dl, "\tilde{\beta}"'] \arrow[d, "\beta"] & \\ 0 \arrow[r] & \mathbf{C} \arrow[r] & C^\infty(M; \mathbf{R}) \arrow[r] & \mathrm{Vect}_{\mathrm{Sp}}(M) \arrow[r] & 0 \end{tikzcd}$$

commutes **[GuSt 1, Kirw, Kost1]**.

REMARKS. (a) A standard basic fact about a semisimple Lie algebra $\mathbf{s}$ is that $H^2(\mathbf{s}; \mathbf{R}) = H^1(\mathbf{s}; \mathbf{R}) = 0$ **[Jaco1]**. Thus if G is semisimple, then any action of G by symplectomorphisms is automatically Hamiltonian, in a unique way.

(b) For a general Lie group G, a symplectic action of G may be regarded as a Hamiltonian action of an appropriate central extension of G; thus the action of a symplectic vector space on itself by translations comes from a Hamiltonian action of the associated Heisenberg group, as in formulas (3.2.1.16)–(3.2.1.19).

Suppose we have a Hamiltonian action β of G on the symplectic manifold M. By duality, the homomorphism $\tilde{\beta}\colon \mathrm{Lie}(G) \to C^\infty(M; \mathbf{R})$ gives us a mapping

$$\mu_\beta\colon M \to \mathbf{g}^*,$$
$$\mu_\beta(m)(x) = \tilde{\beta}(x)(m), \qquad m \in M,\ x \in \mathbf{g}.$$

It is easy to see that the mapping μ_β is equivariant for the action of G. Because μ_β describes the angular momentum of a particle in a particular case (the action of 0_3 on $\mathbf{R}^3 \times \mathbf{R}^{3*} \simeq T^*(\mathbf{R}^3)$ [**AbMa, GuSt1**]), it is called the *moment map.*

The geometry of the moment map for a general Hamiltonian action is quite interesting, and quite relevant for representation theory [**Ati2, GuSt3, Kirw2, DuHV**]. But right now we focus on the case when G acts transitively on M. In this case, the image of μ_β is clearly a single coadjoint orbit. Further, an elementary argument shows that μ_β must be locally a diffeomorphism. Thus any homogeneous Hamiltonian G-action must be a covering space of some coadjoint orbit [**GuSt1, Kost1**]. Or in other words, up to coverings, coadjoint orbits provide the universal examples of transitive Hamiltonian G-actions.

3.2.4. At the start of §3.2 we made a vague reference to the notion of a "classical dynamical system" for G. Now we can specify that we will take this to mean a Hamiltonian G-action. The rationale for this choice comes from the Hamiltonian version of classical mechanics, which shows that a classical conservative dynamical system satisfying Newton's Laws can be expressed as a Hamiltonian action of $\mathbf{R}$ [**AbMa, Arno**]; besides this it has been observed to work.

Given this meaning of "classical dynamical system," the discussion of §3.2.3 can be taken as showing that the irreducible, i.e., transitive, classical dynamical systems for G correspond to coverings of coadjoint orbits. Thus the principle enunciated rather imprecisely at the start of §3.2 can now be stated more clearly: we hope to be able to associate irreducible unitary representations to (covers of) coadjoint orbits for G. The extent to which this hope is realized will be surveyed in the next subsections.

3.3. *Nilpotent groups.* The hope expressed in §3.2.4 is realized perfectly for nilpotent groups, as was discovered by Kirillov [**Kiri, Puka1, Moor**]. (Stating things this way is, in historical terms, to put the cart before the horse; Kirillov's work was a primary inspiration for the philosophy expressed in §3.2.)

3.3.1. A key notion in Kirillov's construction is that of *polarization.* Recall the discussion of coadjoint orbits in §3.2.2. Let G be a Lie group, $\mathbf{g} = \mathrm{Lie}(G)$, $\lambda \in \mathbf{g}^*$, $R_\lambda =$ the stabilizer of λ under Ad^*, and $\mathbf{r}_\lambda = \mathrm{Lie}(R_\lambda)$. By a *polarization* for λ, or *polarizing subalgebra*, or *maximal subordinate subalgebra* we mean a Lie subalgebra $\mathbf{p}$ of $\mathbf{g}$ such that

(3.3.1.1) $\qquad \mathbf{p}$ is a maximal isotropic subspace for the form$\langle\ ,\ \rangle_\lambda$.

Isotropic means that $\langle x, y\rangle_\lambda = 0$ for all $x, y \in \mathbf{p}$. Maximal isotropic of course then means that there is no subspace of $\mathbf{g}$ which properly contains $\mathbf{p}$ and which also is isotropic for $\langle\ ,\ \rangle_\lambda$. The duality theorems of basic linear algebra imply that if $\mathbf{p}$ is a polarization then

$$(3.3.1.2) \qquad \begin{array}{l} \text{(i) } \mathbf{r}_\lambda \subseteq \mathbf{p}, \\ \text{(ii) } \dim \mathbf{p} = \tfrac{1}{2}(\dim \mathbf{r}_\lambda + \dim \mathbf{g}). \end{array}$$

Thus the single condition of (3.3.1.1) could be replaced by the two conditions: (i) $\mathbf{p} \supseteq \mathbf{r}_\lambda$, and (ii) $\mathbf{p}/\mathbf{r}_\lambda$ is maximal isotropic for the symplectic form defined by $\langle\ ,\ \rangle_\lambda$ on $\mathbf{g}/\mathbf{r}_\lambda$.

Before stating Kirillov's results, we should note that for a connected, simply connected nilpotent group N, the exponential map $\exp\colon \mathrm{Lie}(N) \to N$ is a diffeomorphism **[Malc, CoGr, Dixm2]**.

THEOREM 3.3.1.3 (Kirillov). *Let N be a connected and simply connected nilpotent Lie group. Set* $\mathrm{Lie}(N) = \mathbf{n}$.

(a) *There is a natural bijection between the unitary dual $\widehat{N}$ and the set $\mathbf{n}^* / \mathrm{Ad}^* N$ of coadjoint orbits for N.*

(b) *Pick $\lambda \in \mathbf{n}^*$. The representation ρ_λ corresponding to the coadjoint orbit $\mathscr{O}_\lambda$ through λ may be realized as follows. Let $\mathbf{p} \subseteq \mathbf{n}$ be a polarization for λ. (These exist.) Let $P = \exp \mathbf{p}$ be the connected subgroup of N with Lie algebra equal to $\mathbf{p}$. It is a closed subgroup of N. Because $\mathbf{p}$ is isotropic for $\langle\ ,\ \rangle_\lambda$, the formula*

$$(3.3.1.4) \qquad \psi_\lambda(\exp x) = e^{2\pi i \lambda(x)}, \qquad x \in \mathbf{p},$$

defines a unitary character of P. The unitarily induced representation (see §§A.1.14 *and* A.1.16)

$$(3.3.1.5) \qquad 2 - \mathrm{ind}_P^G \psi_\lambda \approx \rho_\lambda$$

is the representation we are looking for.

(c) *Every element of $\widehat{N}$ is strongly trace class (see* §A.1.18). *For an* $\mathrm{Ad}^* N$-*orbit $\mathscr{O} \subseteq \mathbf{n}^*$, the character $\mathscr{O}_{\rho_\mathscr{O}}$ of the corresponding representation $\rho_\mathscr{O}$ can be computed as follows. On $\mathscr{O}$, there is an $\mathrm{Ad}^* N$-invariant measure, unique up to multiples. Denote it by $d_\mathscr{O}\mu$. For $f \in C_c^\infty(N)$, let $f \circ \exp \in C_c^\infty(\mathbf{n})$ be the pullback to $\mathbf{n}$ of f via the exponential map. Define the Fourier transform from functions on $\mathbf{n}$ to functions on $\mathbf{n}^*$ in the usual way:*

$$(3.3.1.6) \qquad \hat{\phi}(\lambda) = \int_{\mathbf{n}} \phi(x) e^{-2\pi i \lambda(x)}\, dx, \qquad \phi \in L^1(\mathbf{n}),$$

where dx is a Haar measure on $\mathbf{n}$. Then for appropriate normalization of the invariant measure $d_\mathscr{O}\mu$ on $\mathscr{O}$, we have

$$(3.3.1.7) \qquad \theta_{\rho_\mathscr{O}}(f) = \int_\mathscr{O} (f \circ \exp)^\wedge(\mu)\, d_\mathscr{O}\mu, \qquad f \in C_c^\infty(N).$$

3.3.2. REMARKS. (a) The proof of Theorem 3.3.1.3 proceeds by induction on the dimension of N, using the tools of the "Mackey Machine" (see **[FeDo, Mack4, Rief]**) for computing representations of group extensions. In fact, the necessary computations are quite limited and depend mainly on understanding the Heisenberg group, the basic group of quantum mechanics, whose Lie algebra is described in formula (3.1.3.1) or (3.2.1.19).

(b) As well as being important for the proof of Theorem 3.3.1.3, the Heisenberg group provides a good illustration of it. Let V be a symplectic vector space. If we again use the isomorphism (3.2.1.1) between V and

V^*, we can write

$$\mathbf{h}(V)^* = (V \oplus \mathbf{R})^* \simeq V \oplus \mathbf{R}. \tag{3.3.2.1}$$

Using the expression (3.2.1.19b) for the Lie bracket in $\mathbf{h}(V)$, and formula (2.4.8) for the adjoint action, we can compute that

(3.3.2.2)

$$\mathrm{Ad}^* \exp(v, t)(v', t') = (v' + t'v, t'), \qquad (v, t) \in \mathbf{h}(V), (v', t') \in \mathbf{h}(V)^*.$$

Denote the connected, simply connected group whose Lie algebra is $\mathbf{h}(V)$ by $H(V)$. From formula (3.3.2.2), we can easily verify the following description of $\mathrm{Ad}^* H(V)$ orbits.

(3.3.2.3) The $\mathrm{Ad}^* H(V)$ orbits in $\mathbf{h}(V)$ are
(i) the points $(v, 0)$, $v' \in V$,
(ii) the hyperplanes $\{(v', t') : v' \in V\}$, $t' \in \mathbf{R} - 0$.

If we plug this data in Theorem 3.3.1.3 we obtain a complete description of the representations of $H(V)$. There are one-dimensional representations $\chi_{v'}(\exp(v, t)) = e^{2\pi i \langle v, v' \rangle}$, $v' \in V$, which factor to the abelian quotient $H(V)/ZH(V)$. Here $ZH(V)$ is the one-dimensional center of $H(V)$; it is also the commutator subgroup. The non-one-dimensional representations correspond to the hyperplanes (3.3.2.3)(ii), and so Theorem 3.3.1.3 specializes to the following classical result (cf. **[CoGr, Foll1, Howe2, Neum]**, etc.).

THEOREM 3.3.2.4 (Stone-von Neumann). *For each nontrivial character χ of $ZH(V)$ $(\simeq \mathbf{R})$, there is up to unitary equivalence exactly one irreducible unitary representation ρ_χ of $H(V)$ with central character χ (see* §A.1.7.4). *The representation ρ_χ may be realized as an induced representation*

$$\rho_X \simeq 2 - \mathrm{ind}_A^H \tilde{\chi}, \tag{3.3.2.5}$$

where $A \subseteq H(V)$ is any maximal abelian connected subgroup, and $\tilde{\chi}$ is any extension of χ from $ZH(V)$ to A.

It is worthwhile to give a concrete description of the representations ρ_χ, to emphasize how close we are here to the heart of classical harmonic analysis **[FePh, Foll1, Howe2, 3]**. For this we can first observe that for $s \in \mathbf{R}^\times$, the map

$$d_s\mathbf{h}(V) \to \mathbf{h}(V), \tag{3.3.2.6}$$
$$d_s(v, t) = (sv, s^2 t), \qquad (v, t) \in \mathbf{h}(V),$$

is an automorphism of $\mathbf{h}(V)$. The corresponding automorphisms of $H(V)$ will permute almost transitively (there will be two orbits which are mutual complex conjugates) the characters of $ZH(V)$. Thus up to the action of the d_s and complex conjugation, there is only one (infinite dimensional unitary irreducible) representation of $H(V)$. So we only need to describe one such representation. But this is in fact given by the realization of $\mathbf{h}(\mathbb{R}^n \oplus (\mathbb{R}^n)^*)$

via the operators $\frac{\partial}{\partial x_j}$ and ix_j on $L^2(\mathbf{R}^n)$, as described in §3.1.3. Thus, via this representation, the universal enveloping algebra $\mathscr{U}(\mathbf{h}(\mathbf{R}^n \oplus \mathbf{R}^{n^*}))$ will be sent to the algebra of polynomial-coefficient differential operators on $\mathbf{R}^n$. (See §A.1.13 for an explanation of how to derive a representation of the enveloping algebra.)

3.4. *Solvable groups.* Kirillov's results appeared in 1960 **[Kiri]**. Through a lot of hard work since then, the basic principles embodied in Kirillov's theory have been extended to encompass a large portion of representation theory of Lie groups. The next class of groups to be analyzed was solvable groups. We briefly outline this development.

Inspection of Theorem 3.3.1.3 makes clear that the bijection of part (a) between orbits and representations is implemented in two quite distinct ways: first, by an explicit construction of the representations, and second by a description of the character of a representation in terms of the orbit. It might seem that a construction of the representation is very much to be preferred to just a description of the character. However, it should be noted that the construction of the representation involves a noncanonical intermediate construction between the orbit and the representation, namely a polarization. While it is always possible to find a polarization, there may in fact be many, and the choice of a particular one is arbitrary. However, one shows that all the representations one constructs by means of various polarizations are equivalent. (A key fact used to do this is the Stone-von Neumann Theorem). This "independence of polarization" allows the construction to succeed. On the other hand, the description of the character via formula (3.3.1.7) is canonical. There is even an a priori description of the proper normalization of the measure $d_{\mathscr{O}}\mu$ **[Moor, Puka2]**. Below we will discuss the generalizations of both parts (b) and (c) to other classes of Lie groups.

3.4.1. For solvable Lie groups, the situation is more complicated, but quite satisfactory. Kirillov's work was generalized almost immediately **[Bert]** to the class known as *exponential solvable groups*, which are characterized as those solvable groups G whose simply-connected cover $\widetilde{G}$ is such that the exponential map exp: $\mathrm{Lie}(\widetilde{G}) \to \widetilde{G}$ is a diffeomorphism **[Moor, Dixm3]**. For exponential solvable groups the bijection between orbits and representations holds, and can be realized using induced representations by an explicit construction using polarizations, just as in the nilpotent case. However, two difficulties arise:

(i) Not all polarizations yield the same representation, or even an irreducible representation;
(ii) Not all representations are strongly trace class.

3.4.1.1. Both difficulties are already illustrated by the two-dimensional "$ax + b$ group"—the group of affine transformations of the line. This may be realized as the set of 2×2 matrices of the form

$$G = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} : b \in \mathbf{R},\ a \in \mathbf{R}^{+\times} \right\}. \tag{3.4.1.1.1}$$

(We restrict a to be positive in order to have a connected group.) The Lie algebra of this group is the space of matrices

$$\begin{bmatrix} \alpha & \beta \\ 0 & 0 \end{bmatrix}, \qquad \alpha, \beta \in \mathbf{R}, \tag{3.4.1.1.2}$$

and its dual may be realized as the space

$$\begin{bmatrix} \lambda & 0 \\ \mu & 0 \end{bmatrix}, \qquad \lambda, \mu \in \mathbf{R}. \tag{3.4.1.1.3}$$

The pairing between the matrices (3.4.1.1.2) and (3.4.1.1.3) is given by taking the trace of products. The coadjoint orbits are

$$\mathscr{O}_\lambda = \left\{ \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \right\} \quad \text{for each } \lambda \in \mathbf{R}, \tag{3.4.1.1.4}$$

$$\mathscr{O}^+ = \left\{ \begin{bmatrix} \lambda & 0 \\ \mu & 0 \end{bmatrix} : \lambda \in \mathbf{R},\ \mu > 0 \right\}, \qquad \mathscr{O}^- = \left\{ \begin{bmatrix} \lambda & 0 \\ \mu & 0 \end{bmatrix} : \lambda \in \mathbf{R},\ \mu < 0 \right\}.$$

The representations corresponding to the one-point orbits are the linear characters (one-dimensional representations) of the group. These are trivial on the commutator subgroup G^1,

$$G^1 = \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} : b \in \mathbf{R} \right\}. \tag{3.4.1.1.5}$$

There are two non-one-dimensional irreducible representations, corresponding to the orbits $\mathscr{O}^+$ and $\mathscr{O}^-$. The Lie algebra of the group G^1 is a polarization for any element in either of these orbits, and we have

$$\rho_{\mathscr{O}^\pm} \simeq 2-\operatorname{ind}_{G^1}^G \chi_\mu, \qquad \begin{bmatrix} 0 \\ \mu \end{bmatrix} \in \mathscr{O}^\pm, \tag{3.4.1.1.6}$$

where $\chi_\mu([\begin{smallmatrix} 1 & b \\ 0 & 0 \end{smallmatrix}]) = e^{2\pi i \mu b}$. However, the group

$$A = \left\{ \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} : a \in \mathbf{R}^{+\times} \right\} \tag{3.4.1.1.7}$$

also defines a polarization of any element of $\mathscr{O}^\pm$. If χ is any character of A, then the unitary representation of G induced from χ is equivalent to the sum $\rho_{\mathscr{O}^+} \oplus \rho_{\mathscr{O}^-}$.

The representations $\rho_{\mathscr{O}^\pm}$ are also not strongly trace class. In fact, if $f \in C_c^\infty(G)$ is such that $\chi_\lambda(f) \neq 0$ for some λ, then $\rho_{\mathscr{O}^\pm}(f)$ will not be trace class. Here χ_λ indicates the character of G corresponding to the orbit $\mathscr{O}_\lambda$ of (3.4.1.1.4). Precisely

$$\chi_\lambda\left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) = a^{2\pi i \lambda}.$$

3.4.1.2. Despite these two complications, the situation for exponential solvable groups is quite well understood. There is a simple criterion first formulated by Pukanszky [**Moor**, **Puka5**] to guarantee that a polarization will produce the appropriate irreducible representation. Further there is a clean description of the representation produced by any polarization [**Moor**, **Vergl**].

With regard to generalizing the trace formula (3.3.1.7), one must recognize that it will not generalize completely because not all representations are strongly trace class. Roughly speaking, it will be closed orbits which correspond to strongly trace class representations. (Observe that the orbits $\mathscr{O}^{\pm}$ of (3.4.1.1.4) are not closed.) Even for orbits for which there is a trace formula, a new phenomenon enters: it is necessary to multiply a function by an appropriate normalizing factor, which depends on the orbit. Thus for an orbit $\mathscr{O}$ for which it is valid, the trace formula takes the form

$$\theta_{\rho_{\mathscr{O}}}(f) = \int_{\mathscr{O}} ((f \circ \exp)(L_{\mathscr{O}}))^{\wedge}(\mu)\, d_{\mathscr{O}}\mu, \qquad f \in C_c^{\infty}(G), \tag{3.4.1.2.1}$$

where $L_{\mathscr{O}}$ is an analytic function on $\mathrm{Lie}(G)$ [**Moor**, **Dufl1**, **Puka4**]. The need to introduce $L_{\mathscr{O}}$ stems from two sources:

(i) For a general exponential solvable group G, the exponential map

$$\exp\colon \mathrm{Lie}(G) \to G$$

will not take Haar measure on $\mathrm{Lie}(G)$ to Haar measure on G.

(ii) The modular function of (the connected subgroup of G whose Lie algebra is) a polarization of $\lambda \in \mathscr{O}$ may not agree with the restriction of the modular function of G.

3.4.2. For general solvable groups, one encounters several difficulties which did not arise in the exponential solvable case.

3.4.2.1. (i) The representations of solvable groups can be badly behaved: these groups need not be type I in the sense of C^*-algebras [**Dixm1**].

(ii) Not all representations are realizable as monomial representations, i.e., as induced representations from linear characters—in terms of the orbit method, this means there are elements $\lambda \in \mathrm{Lie}(G)^*$ for which there is no polarizing subalgebra in $\mathrm{Lie}(G)$. An example is provided below.

(iii) Orbits $\mathscr{O} \subseteq \mathrm{Lie}(G)^*$ may not be simply connected—equivalently, their isotropy groups in $\widetilde{G}$, the simply-connected cover of G, may not be connected (consider E_2, the isometry group of the Euclidean plane). Also, orbit structure may be highly irregular—orbits may not even be locally closed. (A semidirect product $\mathbf{R} \ltimes \mathbf{R}^4$, where $\mathbf{R}$ acts on $\mathbf{R}^4$ by a sum of mutually irrational rotations, provides the simplest example. It was first noted by Mautner.)

These phenomena force a substantial revision in the orbit method, and a naive one-to-one correspondence between coadjoint orbits and representations no longer exists. However, there still exists a highly satisfactory, detailed theory which retains much of the flavor of Theorem 3.3.1.3 [**AuKo**, **Moor**, **Puka3**, **Puka6**]. We will describe how the new features of this theory solve the problems 3.4.2.1.

First one must lump coadjoint orbits into equivalence classes of "quasiorbits." Two orbits define the same quasiorbit if their closures in $\mathrm{Lie}(G)^*$ are equal. It turns out that a quasiorbit in $\mathrm{Lie}(G)^*$ is an orbit for a slightly larger group $G' \supseteq G$ such that the quotient G'/G is abelian [**Puka3**].

Second, one must seek to parametrize, not representations, but primitive ideals in $C^*(G)$ (see [**Dixm2**]). Here $C^*(G)$ is the group C^*-algebra (cf. §A.1.12, especially definition (A.1.12.6)). Recall that if the group G is type I, then there is a natural bijection between equivalence classes of irreducible unitary representations and primitive ideals in $C^*(G)$ [**Dixm2**]. But if G is not type I, there may be infinitely many irreducible representations whose kernel in $C^*(G)$ is a given primitive ideal. (This is in fact fairly typical behavior. The easiest examples may be induced representations of the rank 3 integral Heisenberg group.)

Third, the mapping to quasiorbits from primitive ideals is many-to-one. The fibers are quotients of the duals of subgroups of the component groups of the isotropy groups [**Puka6**].

Let us state the result precisely. Let G be a connected and simply connected solvable Lie group, with Lie algebra $\mathrm{Lie}(G)$. Consider λ in $\mathrm{Lie}(G)^*$. Let $R_\lambda \subseteq G$ be the stabilizer of λ under the coadjoint action Ad^*G, and let R_λ^0 be the identity component of R_λ. Recall that $\mathbf{r}_\lambda = \mathrm{Lie}(R_\lambda^0)$ is the radical of the form $\langle\ ,\ \rangle_\lambda$ associated to λ (cf. (3.2.2.4)). There is a unique character χ_λ on R_λ^0 defined by

$$\chi_\lambda(\exp r) = e^{2\pi i\lambda(r)}, \qquad r \in \mathbf{r}_\lambda. \tag{3.4.2.2}$$

It is easy to see that the component group R_λ/R_λ^0 is abelian. Also, since R_λ stabilizes λ, the character χ_λ on R_λ^0 is clearly invariant under conjugation by all of R_λ. Hence the quotient group $R^\lambda/\ker\chi_\lambda$ is a central extension of R_λ/R_λ^0 by the group $R_\lambda^0/\ker\lambda$, which we may identify to the unit circle $\mathbf{T}$ by means of χ. Thus $R_\lambda/\ker\chi_\lambda$ is a two-step nilpotent group (or possibly abelian—the extension may split), and we have an exact sequence

$$1 \to \mathbf{T} \to R_\lambda/\ker\ \chi_\lambda \to R_\lambda/R_\lambda^0 \to 1. \tag{3.4.2.3}$$

Let S_λ be the image in R_λ/R_λ^0 of the center of $R_\lambda/\ker\chi_\lambda$.

THEOREM 3.4.2.4 [**Puka6**]. *Let G be a connected, simply connected solvable Lie group. Let $\mathbf{P}(\mathbf{C}^*(G)) = \mathbf{P}(G)$ denote the space of primitive ideals of $\mathbf{C}^*(G)$. Let $(\mathrm{Lie}(G)^*/\mathrm{Ad}^*G)^\sim$ be the space of coadjoint quasiorbits. There is* *a mapping*

$$\begin{array}{c} \mathbf{P}(G) \\ \kappa \big\downarrow \\ (\mathrm{Lie}(G)^*/\mathrm{Ad}^* G)^{\sim} \end{array} \tag{3.4.2.5}$$

such that the fiber $\kappa^{-1}(\mathscr{O})$ *above a quasiorbit* $\widetilde{\mathscr{O}} \subseteq \mathrm{Lie}(G^*)$ *can be identified with a quotient of* $\widehat{S}_\lambda$ *for any* $\lambda \in \widetilde{\mathscr{O}}$.

The most subtle aspect of this result is to understand which quotient of $\widehat{S}_\lambda$ gives the fiber. This is closely related to the group G' mentioned above for which the G-quasiorbit becomes an ordinary orbit. If the quasiorbit consists of a single G-orbit, then the fiber is all of $\widehat{S}_\lambda$.

We can also use the notions just formulated to give the criterion of Auslander-Kostant that a solvable group be type I.

THEOREM 3.4.2.6 [**AuKo**]. *The group* G *is type* I *if and only if*

(i) *all coadjoint quasiorbits consist of a single coadjoint orbit, equivalently, the coadjoint orbits are locally closed, and*
(ii) *for every* λ, $S_\lambda = R_\lambda / R_\lambda^0$, *i.e., the extension* (3.4.2.3) *is trivial.*

3.4.3. The correspondence (3.4.2.5) is again described in the two ways indicated by Theorem 3.3.1.3—by explicit construction of induced representations, and by character formulas. However, both these constructions must be more sophisticated. The character formula is similar to the formula (3.4.1.2.1) for exponential groups, except one must restrict the functions f to have support in a certain neighborhood of the identity, and the formula does not distinguish between different elements in the fibers of the map κ of (3.4.2.5) [**Puka3**]. By considering integrals over quasiorbits rather than orbits, Pukanszky [**Puka3**] has formulated an extension of the character formula to the non-type I case.

Although polarizations no longer exist for an arbitrary $\lambda \in \mathrm{Lie}(G)^*$, there is still a fairly direct construction of the representation associated to an orbit as a representation induced from a special class of representations of subgroups. Here again the Heisenberg group, and somewhat more general two-step nilpotent groups, play a key role.

One can preserve the geometric flavor that polarizations give to the constructions by considering *complex polarizations*. In essence, a complex polarization is a complex Lie subalgebra of $\mathrm{Lie}(G)_{\mathbf{C}}$, the complexification of $\mathrm{Lie}(G)$, which satisfies condition (3.3.1.1), where λ now means the complex-linear extension of $\lambda \in \mathrm{Lie}(G)^*$ to $\mathrm{Lie}(G)_{\mathbf{C}}$. In order for a complex polarization in the above sense to be usable for constructing representations, it should also satisfy some other technical conditions [**Moor**, p. 21; **AuKo**], which are usually incorporated into the definition of complex polarization. One can show that complex polarizations always exist. Indeed, Auslander-Kostant establish the existence of complex polarizations satisfying an additional condition called *positivity*. The existence of positive complex polarizations is, once again, essentially a phenomenon associated with the Heisenberg group [**AuKo**, **Moor**].

Having a positive complex polarization for $\lambda \in \mathrm{Lie}(G)^*$ allows one to construct the representation associated to the coadjoint orbit through λ on a space of partially holomorphic sections of a complex line bundle. The basic example is the "Fock model" (cf. [**Barg**, **Foll**, **Howe2**, **Sega**], etc.), for the representations of the Heisenberg group. More recently, several authors [**Carm**, **MoVe**, **Penn**, **Rose**] have considered using nonpositive complex polarizations. This leads to the realization of representations on spaces of higher cohomology of the associated line bundles, rather than sections (= degree zero cohomology). Although these constructions using higher cohomology are not necessary to construct the representations of our solvable G, they establish a parallel between solvable groups and semisimple groups, for which realizations on cohomology are necessary (see §§3.5.5, 3.6.3, 3.6.5).

3.4.4. To conclude our discussion of solvable groups, we will give the basic example showing that polarizations may not exist for all $\lambda \in \mathrm{Lie}(G)^*$, and, correspondingly, that representations of G may not be monomial (i.e., induced from one-dimensional representations of subgroups). The reason not all representations of solvable groups are monomial is related to the age-old fact that not all real matrices are diagonalizable, or even triangularizable, over the real numbers. The four-dimensional Lie algebra described in formula (3.1.1.4) typifies the problem. It may be realized as a Lie algebra of 4×4 matrices:

$$(3.4.4.1) \qquad \mathbf{g} = \left\{ \begin{bmatrix} 0 & x & y & 2z \\ 0 & 0 & -t & y \\ 0 & t & 0 & -x \\ 0 & 0 & 0 & 0 \end{bmatrix} : t, x, y, z \in \mathbf{R} \right\}.$$

The three-dimensional subalgebra of elements of $\mathbf{g}$ with $t = 0$ is a Heisenberg Lie algebra. Denote it by $\mathbf{h}$. The center of $\mathbf{h}$ consists of the elements of $\mathbf{h}$ with $x = 0 = y$. Denote it by $z(\mathbf{h})$. Then $\mathbf{h}/z(\mathbf{h})$ is abelian, and it is easily seen that the adjoint action of $\mathbf{g}/\mathbf{h}$ on $\mathbf{h}/z(\mathbf{h})$ is irreducible (over $\mathbf{R}$—when complexified it will of course break up into a sum of two eigenlines). Consider any λ in $\mathbf{g}^*$ whose restriction to $z(\mathbf{h})$ is nonzero. Simple computations show that the coadjoint orbit $\mathscr{O}_\lambda = \mathscr{O}$ through λ is two-dimensional. Thus the isotropy subalgebra $\mathbf{r}_\lambda$ of λ is also two-dimensional, and any polarization of λ must have dimension 3. However, the projection of $\mathscr{O}_\lambda$ into $\mathbf{h}^*$ is also two-dimensional (see (3.3.2.3)), hence $\dim(\mathbf{r}_\lambda \cap \mathbf{h}) = 1$, so $\mathbf{r}_\lambda$ projects onto $\mathbf{g}/\mathbf{h}$. Since $\mathbf{g}/\mathbf{h}$ acts irreducibly on $\mathbf{h}/z(\mathbf{h})$, there are no three-dimensional subalgebras of $\mathbf{g}$ containing $\mathbf{r}_\lambda$. So λ has no polarizations.

On the other hand, the calculations of §3.1.1 produce a representation of $\mathbf{g}$, acting on the same space as the canonical representation of $\mathbf{h}$, described by the Stone-von Neumann Theorem (Theorem 3.3.2.4). Using this extension of representations from $\mathbf{h}$ to $\mathbf{g}$, one can verify a one-to-one correspondence between orbits and representations for the simply-connected group $\widetilde{G}$ associated to $\mathbf{g}$. Similar, somewhat more general, constructions involving Heisenberg-like groups suffice to construct factor representations corresponding to arbitrary primitive ideals of $\mathbf{C}^*(G)$ for general solvable groups G.

3.5. *Compact groups.* The representation theory of compact Lie groups is equivalent, via the process of differentiating a representation (see §A.1.13), to the representation theory of complex semisimple (actually, reductive) Lie algebras. The bare essence of this is Cartan's theory of the highest weight, and is a key chapter in his foundational work on Lie theory **[Crtn2]**. (For an interesting account of some history of this, see **[Hawk1]**.) Weyl **[Weyl1, PeWe]**, provided the analytic apparatus to make the connection between the two theories, and provided important supplements (complete reducibility, character formula). Harish-Chandra **[HaCh8]** made a connection with the orbit method by providing an orbital interpretation of the Weyl character formula. It is interesting that this work, which is a key to Harish-Chandra's later construction of the discrete series for noncompact semisimple groups, precedes Kirillov's **[Kiri]** by several years. The other aspect of the orbit method, construction of representations via polarizations, is provided by the Borel-Weil-Bott Theorem **[Bott, Warn, Voga 1]**, which is also a development of the 1950s. It too provided important guidance to the noncompact case. In the sections below, we will review these developments more closely.

3.5.1. To start, let us review the representations of $\mathbf{sl}_2$, the unique simple Lie algebra over $\mathbf{C}$ of minimal dimension, namely three. This is a simple and attractive topic, with numerous applications, both within Lie theory proper (cf. §2.8) and in many other parts of mathematics (cf. **[Lang1, HoTa, Howe1,** §4(b); **Proc]**, etc.) and physics (cf. **[BiLo1, 2, Hame, Jone, Shan]**, etc.).

Recall (see formulas (2.8.1)) that $\mathbf{sl}_2$ has a basis h, e^+, e^-, satisfying the commutation relations

$$[h, e^{\pm}] = \pm 2e^{\pm}, \qquad [e^+, e^-] = h. \tag{3.5.1.1}$$

REMARK. We note that the compact group whose complexified Lie algebra is $\mathbf{sl}_2$ is SU_2, the special unitary group in two variables. A basis for $\mathbf{su}_2$ is provided by the famous *Pauli spin matrices* **[Shan]**

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The basis $h, e^{\pm}$ of $\mathbf{sl}_2$ is expressed in terms of the spin matrices as follows:

$$h = \sigma_z, \quad e^+ = \tfrac{1}{2}(\sigma_x + i\sigma_y), \quad e^- = \tfrac{1}{2}(\sigma_x - i\sigma_y).$$

The passage from SU_2 to $\mathbf{su}_2$ to $\mathbf{sl}_2$ is typical of the flexibility permitted by Lie theory.

The first basic fact about representations of $\mathbf{sl}_2$ stems directly from the first of these relations. Suppose we have a triple of operators $h, e^{\pm}$ on a vector space V, and suppose $h, e^{\pm}$ satisfy the commutation relations (3.5.1.1). Suppose v is an eigenvector for h, with eigenvector λ:

$$h(v) = \lambda v. \tag{3.5.1.2}$$

We compute

$$\begin{aligned} h(e^{+}(v)) &= ([h, e^{+}] + e^{+}h)(v) = 2e^{+}(v) + e^{+}(\lambda v) \\ &= (\lambda + 2)e^{+}(v). \end{aligned} \tag{3.5.1.3}$$

Thus $e^{+}(v)$ is again an eigenvector for h, with eigenvalue $\lambda + 2$, the eigenvalue of v plus two. A similar computation shows $e^{-}(v)$ is also an h-eigenvector, with eigenvalue $\lambda - 2$. Thus the effect of e^{+} is to shift the eigenspaces of h to higher eigenvalues; e^{-} shifts the h-eigenspaces toward lower eigenvalues. This phenomenon is commonly described by calling e^{+} a *raising operator* and e^{-} a *lowering operator.* We may summarize the above computation as follows.

LEMMA 3.5.1.4. *If V is a module for $\mathbf{sl}_2$, and $V_\lambda \subseteq V$ is the λ-eigenspace for h, then the sum $\sum_{k\in\mathbf{Z}} V_{\lambda+2k}$ of h-eigenspaces is invariant under $\mathbf{sl}_2$. More precisely, we have*

$$e^{\pm}(V_{\lambda+2k}) \subseteq V_{\lambda+2(k\pm1)}.$$

The above discussion shows that the product $e^{-}e^{+}$ preserves h-eigenspaces. For a sharper understanding of the structure of representations of $\mathbf{sl}_2$, we investigate the structure of the operator $e^{-}e^{+}$ (or $e^{+}e^{-}$; but $e^{+}e^{-} = e^{-}e^{+} + h$).

To analyze $e^{-}e^{+}$ we consider the *Casimir operator*

$$\mathscr{C} = h^2 + 2(e^{+}e^{-} + e^{-}e^{+}) = h^2 + 2h + 4e^{-}e^{+} = h^2 - 2h + 4e^{+}e^{-}. \tag{3.5.1.5}$$

A straightforward computation shows that $\mathscr{C}$ commutes with all of $\mathbf{sl}_2$. Thus $\mathscr{C}$ is in the center of the universal enveloping algebra of $\mathbf{sl}_2$. In fact, it generates the center (cf. **[Lang1, Hump]**, etc.).

Since $\mathscr{C}$ commutes with $\mathbf{sl}_2$, its eigenspaces will be invariant under $\mathbf{sl}_2$. If V consists of a single eigenspace for $\mathscr{C}$, we will say the action of $\mathbf{sl}_2$ on V is *quasisimple.* Clearly all finite-dimensional irreducible representations are quasisimple by Schur's Lemma (cf. **[HeRo, Lang3, Jaco2, Knap2]** etc.).

Suppose the action of $\mathbf{sl}_2$ on V is quasisimple, so that $\mathscr{C}$ acts on V by a scalar, which we will denote by μ. Again let V_λ be the λ-eigenspace for h. Then if $v \in V_\lambda$, equation (3.5.1.5) says

$$e^{-}e^{+}(v) = \tfrac{1}{4}(\mu - \lambda^2 - 2\lambda)v, \qquad e^{+}e^{-}(v) = \tfrac{1}{4}(\mu - \lambda^2 + 2\lambda)(v). \tag{3.5.1.6}$$

Thus, if V is quasisimple, the operator $e^{-}e^{+}$ acts as a scalar on each V_λ, and this scalar will be nonzero unless we have the quadratic relation

$$\mu + 1 = (\lambda + 1)^2. \tag{3.5.1.7a}$$

Similarly, e^+e^- acts as a scalar, which is nonzero unless

$$\mu + 1 = (\lambda - 1)^2. \tag{3.5.1.7b}$$

Note that (3.5.1.7a) becomes (3.1.5.7b) under the translation $\lambda \to \lambda + 2$. These equations imply that in a sum $\sum_{k\in\mathbf{Z}} V_{\lambda+2k}$, there are at most two values of k for which either of the maps

$$e^+ : V_{\lambda+2k} \to V_{\lambda+2k+2}, \qquad e^- : V_{\lambda+2k+2} \to V_{\lambda+2k}$$

fails to be an isomorphism.

Now suppose V is irreducible and finite dimensional. Then necessarily $V = \sum_{k\in\mathbf{Z}} V_{\lambda+2k}$ for some fixed λ, and clearly $V_{\lambda+2k} = \{0\}$ for k large enough. By replacing λ by $\lambda + 2k$ for appropriate k, we can assume $V_\lambda \neq \{0\}$ but $V_{\lambda+2k} = \{0\}$ for $k > 0$. Choose $v_0 \in V_\lambda$, and set $v_j = (e^-)^j\, v_j$. Then we have the formulas

$$e^-(v_j) = v_{j+1}, \qquad e^+v_j = j(\lambda - j + 1)v_{j-1}. \tag{3.5.1.8}$$

The first formula amounts to the definition of v_j, and the second follows from formulas (3.5.1.6). Since V is finite dimensional, we must have $e^-(v_j) = 0$ for some j. Then also $e^+e^-(v_j) = e^+(v_{j+1}) = 0$. From the second of formulas (3.5.1.8), we see that necessarily $j = \lambda$. Hence λ must be a nonnegative integer. We then further see that formulas (3.5.1.8) define a unique $\mathbf{sl}_2$-module structure on the span of the v_j, $0 \leq j \leq \lambda$. This span must thus be all of V. We conclude $\dim V = \lambda + 1$, and that the v_j's are a basis for V. The following result summarizes our analysis.

PROPOSITION 3.5.1.9. *For each positive integer n, there is up to isomorphism a unique irreducible representation of $\mathbf{sl}_2$ of dimension n. The space of this representation allows a basis $\{v_j : 0 \leq j \leq n-1\}$ with respect to which the action of $\mathbf{sl}_2$ is described by* (3.5.1.8), *with $\lambda = n - 1$. In particular, the eigenvalues of h are*

$$\{m : -(n-1) \leq m \leq n-1,\ m \equiv n-1 \pmod 2\}$$

and these eigenvalues all have multiplicity one.

REMARK. The above arguments can easily be adapted to describe all irreducible representations, finite- or infinite-dimensional, of the group $SL_2(\mathbf{R})$ (cf. **[Barg1, Lang1, HoTa]**, etc.).

3.5.2. The very precise picture presented by Proposition 3.5.1.9 has an analog for a general (semi-) simple complex Lie algebra $\mathbf{g}$. The basic results are due to Cartan **[Crtn2]** but understanding of the structure behind them has been refined considerably since 1913. We will give a fairly modern account, based roughly on **[HaCh1, Jaco1, Hump, BGG1–3]**.

To set the mood for this construction, consider the following description of the finite-dimensional representations of $\mathbf{sl}_2$. If V is an $\mathbf{sl}_2$-module and

$v \in V$, call v a *highest weight vector* if $e^+(v) = 0$ and $h(v) = \lambda v$ for some number λ. The eigenvalue of λ is then called a *highest weight.* The vector v_0 of the basis v_j in (3.5.1.8) is a highest weight vector with highest weight λ. An $\mathbf{sl}_2$-module which is generated by a highest weight vector is called a *highest weight module.* It is easy to see that if we have a vector space $V(\lambda)$ with basis $\{v_j : 0 \leq j < \infty\}$, and we define an action of $\mathbf{sl}_2$ on $V(\lambda)$ by formulas (3.5.1.8), then we obtain a highest weight module, with v_0 a highest weight vector of weight λ. Further, an easy argument, again based on formulas (3.5.1.6), shows $V(\lambda)$ is the universal highest weight module with highest weight λ in the sense that if $U(\lambda)$ is any highest weight with highest weight λ, there is a surjective $\mathbf{sl}_2$-module morphism from $V(\lambda)$ to $U(\lambda)$.

We can do this for any number λ. Typically $V(\lambda)$ is irreducible. However, if λ is a nonnegative integer, then the quantity $j(\lambda - j + 1)$ will be zero not only for $j = 0$, but also when $j = \lambda + 1$. In this case, $v_{\lambda+1}$ will be a highest weight vector, with highest weight $-\lambda - 2$. Thus, when λ is a nonnegative integer, the module $V(-\lambda - 2)$ is a submodule of $V(\lambda)$. One sees that the quotient $V(\lambda)/V(-\lambda-2)$ is the finite-dimensional irreducible representation of dimension $\lambda + 1$. Another way of saying this is to observe that we have an exact sequence

$$(3.5.2.1) \qquad 0 \to V(-\lambda - 2) \to V(\lambda) \to F(\lambda) \to 0,$$

where $F(\lambda)$ is the finite-dimensional irreducible representation with highest weight λ.

3.5.3. The description (3.5.2.1) of the finite-dimensional irreducible representation of $\mathbf{sl}_2$ has a generalization to all complex semisimple (finite-dimensional) Lie algebras [**Jaco1, Hump, BGG1–3**]. We will describe it. Let $\mathbf{g}$ be a complex semisimple Lie algebra, and let $\mathbf{a} \subseteq \mathbf{g}$ be a Cartan subalgebra. Consider the decomposition (2.8.6) of $\mathbf{g}$ into root spaces for $\mathbf{a}$:

$$\mathbf{g} = \mathbf{a} \oplus \sum_{\alpha \in \Sigma} \mathbf{g}_\alpha.$$

Here Σ denotes the set of roots of $\mathbf{a}$ acting on $\mathbf{g}$. As we remarked in §2.8, if $\mathbf{g}_\alpha$ is a root space, then so is $\mathbf{g}_{-\alpha}$, and $\mathbf{g}_\alpha$ and $\mathbf{g}_{-\alpha}$ together generate an algebra $\mathbf{s}_\alpha$ isomorphic to $\mathbf{sl}_2$. Let h_α in $\mathbf{s}_\alpha \cap \mathbf{a}$ be the element corresponding to the element h as in formulas (2.8.1). In other words, the element h_α is determined by the conditions

$$(3.5.3.1) \qquad h_\alpha \in \mathbf{a} \cap \mathbf{s}_\alpha, \qquad \alpha(h_\alpha) = 2.$$

The element h_α is frequently called a *coroot.*

It follows from the description in Proposition 3.5.1.9 of the representations of $\mathbf{sl}_2$ that $\beta(h_\alpha) \in \mathbf{Z}$ for all roots β. Thus if we denote by $\mathbf{a}_\mathbf{R}$ the real span of the h_α for all roots α, we see that the roots take real values on $\mathbf{a}_\mathbf{R}$.

Choose any $h_0 \in \mathbf{a}_{\mathbf{R}}$ such that $\alpha(h_0) \neq 0$ for all $\alpha \in \Sigma$. Set

(3.5.3.2) $$\Sigma^+ = \{\alpha \in \Sigma : \alpha(h_0) > 0\}, \qquad \Sigma^- = -\Sigma^+ = \{\alpha : \alpha(h_0) < 0\}.$$

The sets Σ^+ and Σ^- are called, respectively, the *positive roots* and the *negative roots.* Further, set

(3.5.3.3) $$\mathbf{n}^+ = \sum_{\alpha \in \Sigma^+} \mathbf{g}_\alpha, \qquad \mathbf{n}^- = \sum_{\alpha \in \Sigma^-} \mathbf{g}_\alpha.$$

Then $\mathbf{n}^+$ and $\mathbf{n}^-$ are maximal nilpotent Lie subalgebras of $\mathbf{g}$, and we have the decomposition

(3.5.3.4) $$\mathbf{g} = \mathbf{a} \oplus \mathbf{n}^+ \oplus \mathbf{n}^-.$$

Further, the algebras

(3.5.3.5) $$\mathbf{b}^+ = \mathbf{a} \oplus \mathbf{n}^+, \qquad \mathbf{b}^- = \mathbf{a} \oplus \mathbf{n}^-$$

are maximal solvable subalgebras of $\mathbf{g}$. They are called *Borel subalgebras.* The commutator subalgebra of $\mathbf{b}^\pm$ is $\mathbf{n}^\pm$.

Suppose we have a representation of $\mathbf{g}$ on a vector space V. Since the algebra $\mathbf{a}$ is commutative, it is possible to have simultaneous eigenvectors for $\mathbf{a}$ in V. Suppose v is such a vector, i.e., suppose that for all a in $\mathbf{a}$ we have $a(v) = \lambda(a)v$ for some number $\lambda(a)$. It is trivial to check that the function

(3.5.3.6) $$\lambda : a \to \lambda(a)$$

depends linearly on a, so that λ belongs to $\mathbf{a}^*$. The linear functional λ is called the *weight* of v, and v is called a *weight vector of weight* λ. The span of all weight vectors of weight λ is called the λ *weight space.* Suppose v is not just an eigenvector for $\mathbf{a}$, but for all of $\mathbf{b}^+$; that is, suppose v is a weight vector for $\mathbf{a}$, and additionally $n(v) = 0$ for $n \in \mathbf{n}^+$. Then v is called a *highest weight vector,* and the weight λ of v is a *highest weight.* If V is generated as a $\mathbf{g}$-module by a highest weight vector, then V is called a *highest weight module.* Just as for $\mathbf{sl}_2$ we can prove

LEMMA 3.5.3.7. *Every finite-dimensional irreducible representation of* $\mathbf{g}$ *is a highest weight module. More precisely, a finite-dimensional irreducible representation contains a unique highest weight vector.*

PROOF. Let V be the space of the representation. Since V is irreducible, to show it is a highest weight module it suffices to show it contains a highest weight vector. This is done in completely elementary fashion just as for $\mathbf{sl}_2$. If $h_0 \in \mathbf{a}$ is the element used to construct $\mathbf{n}^\pm$, observe that if $n \in \mathbf{g}_\alpha \subseteq \mathbf{n}^+$, then n transforms an eigenvector for h_0 of eigenvalue λ into an eigenvector of eigenvalue $\lambda + \alpha(h_0)$, which has larger real part than does λ. Hence if λ has maximal real part among the eigenvalues of h_0 acting on V, then any eigenvector for h_0 with eigenvalue λ must be annihilated by $\mathbf{n}^+$. Since

$\mathbf{a}$ is commutative and V is finite dimensional, we may find within the λ-eigenspace for h_0 a weight vector for $\mathbf{a}$. It is necessarily then a highest weight vector.

To show there is only one highest weight vector, we appeal to the Poincaré-Birkhoff-Witt Theorem (cf. [**Jaco1**, **Serr2**], etc.). From equations (3.5.3.4) and (3.5.3.5) we see that

$$\mathbf{g} = \mathbf{b}^+ \oplus \mathbf{n}^- .$$

Let $\mathscr{U}(\mathbf{g})$ be the universal enveloping algebra of $\mathbf{g}$ (cf. [**Jaco1**, **Serr2**], etc.), and similarly for $\mathbf{b}^+$, $\mathbf{n}^-$. Multiplication inside $\mathscr{U}(\mathbf{g})$ induces a linear mapping

$$\mathscr{U}(\mathbf{n}^-) \otimes \mathscr{U}(\mathbf{b}^+) \to \mathscr{U}(\mathbf{g}). \tag{3.5.3.8}$$

The PBW Theorem tells us that the mapping (3.5.3.8) is a linear isomorphism.

Let $v \in V$ be a highest weight vector. Denote by $\mathbf{C}v$ the line through v. Then using PBW we find

$$\mathscr{U}(\mathbf{g})(\mathbf{C}v) = \mathscr{U}(\mathbf{n}^-)\mathscr{U}(\mathbf{b}^+)(\mathbf{C}v) = \mathscr{U}(\mathbf{n}^-)(\mathbf{C}v).$$

For each $\alpha \in \Sigma^-$, choose a nonzero element $n_\alpha \in \mathbf{g}_\alpha$. Then $\mathscr{U}(\mathbf{n}^-)$ is spanned by monomials in the n_α, i.e., by products $n_{\alpha_1} n_{\alpha_2} \cdots n_{\alpha_k}$. An easy inductive calculation shows that, if v has weight λ, then $n_{\alpha_1} \cdots n_{\alpha_k}(v)$ is also a weight vector, of weight $\lambda + \sum_{i=1}^k \alpha_i$. We note that since $\alpha(h_0) < 0$ for all α in Σ^-, no sum $\sum_{i=1}^k \alpha_i$ can be zero unless $k = 0$. Thus we have the following result.

LEMMA 3.5.3.9. *If V is a highest weight module with highest weight λ, then:*

(i) *V is a direct sum of its weight spaces;*
(ii) *all weights of V have the form $\lambda + \sum_{\alpha \in \Sigma^-} n_\alpha \alpha$, where the n_α are nonnegative integers; and*
(iii) *the λ-weight space is one-dimensional, that is, it is $\mathbf{C}v$, where v is the highest weight vector of weight λ.*

Now suppose V is an irreducible highest weight module, with highest weight λ, and suppose V contains a highest weight vector v_1 in addition to the highest weight vector v of weight λ. Then by Lemma 3.5.3.9(ii) and (iii), the weight of v_1 is $\lambda + \sum_{\alpha \in \Sigma^-} m_\alpha \alpha$, with some of the m_α's positive. By Lemma 3.5.3.9(ii) the $\mathbf{g}$-module $\mathscr{U}(\mathbf{g})(v_1)$ is the span of weight spaces with weights $\lambda + \sum_{\alpha \in \Sigma^-}(m_\alpha + n_\alpha)\alpha$, with the n_α's nonnegative. It follows that v cannot belong to $\mathscr{U}(\mathbf{g})(v_1)$, contradicting the irreducibility of V. This proves Lemma 3.5.3.7.

In fact, during the argument, we showed a more general fact about highest weight modules, which we will state explicitly.

COROLLARY 3.5.3.10. *Let V be a highest weight module, generated by the highest weight vector v with highest weight λ. Then*

(i) *V is irreducible if and only if V contains no other highest weight vector; and*

(ii) *V contains a unique maximal proper submodule U such that V/U is irreducible and nontrivial. (In particular, $v \notin U$.) U is generated by all highest weight vectors other than v.*

Thus we have identified irreducible finite-dimensional representations as members of a larger family of irreducible highest weight modules. We will now proceed by describing this larger class, then identifying the subclass consisting of finite-dimensional subrepresentations.

First, we show that, as for $\mathbf{sl}_2$, there is a highest weight module with highest weight λ for any $\lambda \in \mathbf{a}^*$. Indeed, given $\lambda \in \mathbf{a}^*$, consider the left ideal $\mathscr{L}_\lambda$ in $\mathscr{U}(\mathbf{g})$ generated by $\mathbf{n}^+$ and by elements $a-\lambda(a)$ for $a \in \mathbf{a}$. Note that λ defines a character (a one-dimensional representation) of $\mathscr{U}(\mathbf{b}^+)$, and that $\mathbf{n}^+$ and the elements $a - \lambda(a)$ generate the kernel of the corresponding homomorphism from $\mathscr{U}(\mathbf{b}^+)$ to $\mathbf{C}$. Thus they generate a two-sided ideal $\mathscr{J}_\lambda$ of codimension one in $\mathscr{U}(\mathbf{b}^+)$, and $\mathscr{L}_\lambda$ is the left ideal in $\mathscr{U}(\mathbf{g})$ generated by $\mathscr{J}_\lambda$. It follows from PBW that $\mathscr{L}_\lambda = \mathscr{U}(\mathbf{n}^-)\mathscr{J}_\lambda$, and that the natural map

$$\mathscr{U}(\mathbf{n}^-) \hookrightarrow \mathscr{U}(\mathbf{g}) \to \mathscr{U}(\mathbf{g})/\mathscr{L}_\lambda \tag{3.5.3.11}$$

is a linear isomorphism.

COROLLARY 3.5.3.12. (a) *The* $\mathbf{g}$*-module*

$$V_\lambda = \mathscr{U}(\mathbf{g})/\mathscr{L}_\lambda \tag{3.5.3.13}$$

is a highest weight module, with highest weight λ, generated by the image v_λ of 1, the identity element of $\mathscr{U}(\mathbf{g})$.

(b) *V_λ is free as a $\mathscr{U}(\mathbf{n}^-)$ module.*

(c) *Any highest weight module with highest weight λ is a quotient of V_λ.*

(d) *Consequently, for every weight $\lambda \in \mathbf{a}^*$, there exists a unique irreducible highest weight module M_λ with highest weight λ.*

The modules V_λ are usually called *Verma modules* **[Hump, BGG3]**.

Thus we have an irreducible highest weight module M_λ for every $\lambda \in \mathbf{a}^*$. It remains to decide when M_λ is finite dimensional. We can deduce some restrictions on λ from our knowledge of $\mathbf{sl}_2$. Suppose M_λ is finite dimensional. For a positive root $\alpha \in \Sigma^+$, consider the copy $\mathbf{s}_\alpha$ of $\mathbf{sl}_2$ generated by $\mathbf{g}_{\pm\alpha}$. The highest weight vector v_λ of M_λ generates a highest weight module for $\mathbf{s}_\alpha$, and this highest weight module is necessarily finite dimensional. It follows from §3.5.2 that $\lambda(h_\alpha)$ is a nonnegative integer. Let us say $\lambda \in \mathbf{a}^*$ is *integral* if $\lambda(h_\alpha)$ is an integer for all $\alpha \in \Sigma^+$. Let us say $\lambda \in \mathbf{a}^*$ is *dominant* if $\lambda(h_\alpha) \geq 0$ for all $\alpha \in \Sigma^+$. (This is equivalent to saying λ is in the positive or fundamental Weyl chamber, cf. §2.10.) Then, for $\lambda \in \mathbf{a}^*$ to be the highest weight of an irreducible finite-dimensional representation of $\mathbf{g}$, we can say it must be dominant and integral. The main result of Cartan's highest weight theory is that these conditions on λ suffice to guarantee M_λ is finite dimensional.

THEOREM 3.5.3.14. *The irreducible module M_λ of highest weight λ is finite dimensional if and only if λ is dominant and integral.*

REMARKS. (a) This theorem reminds us again of the strong control $\mathbf{sl}_2$ exerts over the phenomena of semisimple Lie algebras. This control is evident even more in the proof of the theorem given below.

(b) The dominant integral λ in $\mathbf{a}^*$ clearly forms a semigroup under addition—the intersection of a lattice with a cone. If M_λ and M_μ are irreducible highest weight modules with highest weight vectors v_λ, v_μ, then the tensor product $v_\lambda \otimes v_\mu$ will generate a highest weight module, of highest weight $\lambda+\mu$, inside the tensor product module $M_\lambda \otimes M_\mu$. Hence, if M_λ, M_μ are finite dimensional, so must $M_{\lambda+\mu}$ be. Thus the set of highest weights of finite-dimensional representations is also a semigroup. To prove that all dominant integral λ define finite-dimensional highest weight modules, it suffices to exhibit finite-dimensional M_λ for a set of λ which generate the semigroup of dominant integral weights. This is essentially what Cartan did **[Crtn2]**, and in fact the procedure, though heavily computational for the exceptional groups, is illuminating, and for the classical groups is quite elegant, involving the exterior powers of the standard representations. From general structure theory **[Jaco1, Hump]** one can show that the dominant integral weights actually form a free semigroup on a unique set of $\operatorname{rank}(\mathbf{g}) = \dim \mathbf{a}$ generators. The representations corresponding to these generators are called the *fundamental representations* of $\mathbf{g}$. For $\mathbf{g} = \mathbf{sl}_n$, the fundamental representations are just the natural action on the $\Lambda^j(\mathbf{C}^n)$, the exterior powers of $\mathbf{C}^n$, for $1 \le j \le n-1$. For orthogonal and symplectic Lie algebras, the fundamental representations (except for the spin representations of the orthogonal algebras **[Arti, BeTu, Jaco2]**) are also constructed fairly easily from the exterior powers of the basic representation.

We will briefly sketch the approach of **[HaCh1]** (see also **[Jaco1, Hump]**) to showing that, if λ is dominant integral, then M_λ is finite dimensional. Consider the fundamental positive roots in Σ^+ (cf. §2.12). Let α be a fundamental positive root. From the general structure theory, we know that $\mathbf{p}_\alpha$, defined by

$$\mathbf{p}_\alpha = \mathbf{b}^+ + \mathbf{g}_{-\alpha} = \mathbf{n}^+_{(\alpha)} \oplus \ker\alpha \oplus \mathbf{s}_\alpha, \tag{3.5.3.15}$$

where $\mathbf{n}^+_{(\alpha)} = \sum_{\beta\in\Sigma^+,\,\beta\neq\alpha} \mathbf{g}_\beta$ and $\ker\alpha = \{h : \alpha(h) = 0\} \subseteq \mathbf{a}$, is a Lie subalgebra of $\mathbf{g}$. It is called a *parabolic subalgebra*. The subspace $\mathbf{n}^+_{(\alpha)}$ is an ideal in $\mathbf{p}_\alpha$. In particular, we have

$$[\mathbf{n}^+_{(\alpha)}, \mathbf{s}_\alpha] \subseteq \mathbf{n}^+_\alpha. \tag{3.5.3.16}$$

Consider the Verma module V_λ with highest weight vector v_λ. Suppose that $\lambda(h_\alpha)$ is a nonnegative integer. Then formulas (3.5.1.8) show that if $e_{-\alpha}$ belongs to $\mathbf{g}_{-\alpha}$, and e_α belongs to $\mathbf{g}_\alpha$, the vector

$$y = e_{-\alpha}^{\lambda(h_\alpha)+1}(v_\lambda)$$

is annihilated by e_α. Also, the commutation relations (3.5.3.16) imply that y will be annihilated by $\mathbf{n}^+_{(\alpha)}$. Since $\mathbf{n}^+ = \mathbf{n}^+_{(\alpha)} \oplus \mathbf{g}_\alpha$, it follows that y is a highest weight vector, of weight $\lambda - (\lambda(h_\alpha)+1)\alpha$. It will generate a highest weight submodule of V_λ. Since $\mathbf{n}^-$ acts freely on V_λ, we see that y generates a module isomorphic to $V_{\lambda-(\lambda(h_\alpha)+1)\alpha}$. In other words, under the hypothesis that $\lambda(h_\alpha)$ is a nonnegative integer, we obtain an embedding of $V_{\lambda-(\lambda(h_\alpha)+1)\alpha}$ in V_λ.

If λ is dominant integral, then we get an embedding of $V_{\lambda-(\lambda(h_\alpha)+1)\alpha}$ in V_λ for every fundamental root α. This already suffices to show that M_λ, the irreducible quotient of V_λ, must be finite dimensional. Indeed, for each fundamental root α, M_λ will be a quotient of $V_\lambda/V_{\lambda-(\lambda(h_\alpha)+1)\alpha} = V_\lambda(\alpha)$. The image in $V_\lambda(\alpha)$ of the highest weight vector v_λ generates a finite-dimensional $\mathbf{s}_\alpha$ module. Since the adjoint action of $\mathbf{g}$ on $\mathscr{U}(\mathbf{g})$ is a sum of finite-dimensional $\mathbf{g}$-modules, hence $\mathbf{s}_\alpha$-modules, it follows that any element of $V_\lambda(\alpha) = \mathscr{U}(\mathbf{g})(v_\lambda)$ generates a finite-dimensional $\mathbf{s}_\alpha$-module. It follows that $S_\alpha = \exp \mathbf{s}_\alpha$, the *group* obtained by exponentiating $\mathbf{s}_\alpha$, acts on $V_\lambda(\alpha)$. In particular, the Weyl group reflection w_α contained in S_α acts on $V_\lambda(\alpha)$. It is easy to see this fact remains true in any quotient $\mathbf{g}$-module of $V_\lambda(\alpha)$. In particular, w_α acts on M_λ. Since the w_α generate the full Weyl group W (cf. §2.9), we see that W acts on M_λ.

Since W normalizes $\mathbf{a}$, the effect of W on M_λ is to permute weight spaces. Precisely, for $\mu \in \mathbf{a}^*$, let M_λ^μ denote the μ weight space of M_λ. Then for $p \in W$, we have

$$p(M_\lambda^\mu) = M_\lambda^{p(\mu)}, \tag{3.5.3.17}$$

where $p(\mu)$ denotes the standard action of p on μ as an element of $\mathbf{a}^*$. Thus, in particular, one sees that the set of weights μ for which $M_\lambda^\mu \neq \{0\}$ is invariant under W. Since also the weights of M_λ, being contained in the weights of V_λ, are bounded above, as described by Lemma 3.5.3.9(ii), it follows easily from the geometry of the action of W on $\mathbf{a}^*$ that the set of weights μ for which $M_\lambda^\mu \neq 0$ must be bounded, hence finite in number. Since each weight space of M_λ (indeed, of V_λ) is finite dimensional, we conclude M_λ is finite dimensional.

Although the argument above gives us the desired finite dimensionality of M_λ when λ is dominant integral, it does not give us a very precise picture of M_λ. A refinement of the above considerations yields a description of M_λ analogous to (3.5.2.1) **[BGG1–3, Dixm1]**.

Let ρ denote the element of $\mathbf{a}^*$ such that

$$\rho(h_\alpha) = 1 \tag{3.5.3.18}$$

for each fundamental root α. Then we may write

$$\lambda - (\lambda(h_\alpha) + 1)\alpha = w_\alpha(\lambda + \rho) - \rho, \tag{3.5.3.19}$$

where again w_α is the Weyl group reflection corresponding to the fundamental root α.

As we have noted, the Weyl group W is generated by the reflections w_α. Let the *length* of $p \in W$ be the shortest product of the w_α's equaling p **[Hill, Bour]**. Denote it by $l(p)$. If

$$p = w_{\alpha_l} w_{\alpha_{l-1}} \cdots w_{\alpha_1}, \qquad l = l(p),$$

is a shortest possible product expressing p, then

$$p' = w_{\alpha_l} p = w_{\alpha_{l-1}} \cdots w_{\alpha_1}$$

has length $l - 1$. From a systematic study of the geometry of a root system and its Weyl group, one can see that if λ is dominant, then $p'(\lambda)(h_{\alpha_l}) \geq 0$. It follows by the argument given above that $V_{p(\lambda+\rho)-\rho}$ embeds in $V_{p'(\lambda+\rho)-\rho}$. The embedding is unique up to multiples.

By induction, we find that when λ is dominant integral, we can embed $V_{w(\lambda+\rho)-\rho}$ in V_λ for every element w of the Weyl group W. It is shown in **[BGG2]** that these embeddings can be organized into an exact sequence, as follows. For $k \geq 0$, set

$$V_\lambda^{(k)} = \sum_{l(w)=k} V_{w(\lambda+\rho)-\rho}. \tag{3.5.3.20}$$

We have seen that whenever w has length $k - 1$ and $w_\alpha w$ has length k, there is an embedding $V_{w_\alpha w(\lambda+\rho)-\rho} \to V_{w(\lambda+\rho)-\rho}$, defined up to multiples. By taking linear combinations of these embeddings, we can construct mappings from $V_\lambda^{(k)}$ to $V_\lambda^{(k-1)}$. If we choose these mappings correctly, we will get an exact sequence

$$0 \to V_\lambda^{(m)} \to V_\lambda^{(m-1)} \to \cdots \to V_\lambda^{(2)} \to V_\lambda^{(1)} \to V_\lambda \to M_\lambda \to 0, \tag{3.5.3.21}$$

where m is the largest possible length of an element of W. In fact, $m = \dim \mathbf{n}^-$. In the case of $\mathbf{sl}_2$, this exact sequence is simply the sequence (3.5.2.1).

REMARKS. (a) The exact sequence (3.5.3.21) implies the Weyl character formula (cf. §3.5.4) by means of the Euler-Poincaré principle. The alternating sum of the highest weights of the $V_{w(\lambda+\rho)-\rho}$ provides the numerator for the formula, while the character of V_0 ($\simeq \mathbf{n}^-$ as an $\mathbf{a}$-module) provides the celebrated "Weyl denominator."

(b) The multiplicities of the weight spaces of $V_0 \simeq \mathbf{n}^-$ are easily seen by PBW to be given by the *Kostant partition function* **[Kost5, Jaco1, Hump]**,

$P(\lambda) = \#$ of ways of expressing λ as an integral linear combination of negative roots. Given this observation, Kostant's multiplicity formula **[Kost5, Jaco1, Hump]** for the multiplicities of weights of finite-dimensional representations follows immediately from (3.5.3.21). Indeed, Kostant's formula is basically a variant way of expressing the Weyl character formula, so when we can deduce one, we should be able to deduce the other.

(c) Also from the exact sequence (3.5.3.21), one can fairly directly deduce Kostant's description **[Kost4, Warn, Knap1, Voga2, Arib]** of the Lie algebra cohomology groups $H^k(\mathbf{n}^+, M_\lambda)$-cohomology of $\mathbf{n}^+$ with coefficients in the module M_λ. We will discuss this in §3.5.5.

(d) The "ρ-shift" seen in the highest weights of the $V_{w(\lambda+\rho)-\rho}$, and in the Weyl character formula, and elsewhere is in some sense explained by the Harish-Chandra homomorphism (cf. Theorem 3.5.5.23)).

3.5.4. Weyl's character formula. In **[Weyl1]**, (see also **[Weyl2, Wall2, Knap]**), Hermann Weyl gave a radically different approach to the representation theory of complex semisimple Lie algebras through the equivalent theory of representations of compact semisimple groups. (Part of his achievement was to make explicit the equivalence. This is the origin of the celebrated "unitary trick.") This approach yields not only the classification of irreducible representations but also a formula for their characters, the Weyl character formula. (We note that the Weyl character formula for U_n (and also for O_n) is due to Schur **[Schu]**.)

We will illustrate the method with the unitary group U_n in order not to become too involved with the notation necessary for the general case.

We think of U_n as a set of $n \times n$ matrices. The subgroup (a Cartan subgroup)

$$(3.5.4.1) \qquad A = \left\{ \begin{bmatrix} z_1 & & & \\ & z_2 & 0 & \\ & & \ddots & \\ & 0 & & z_n \end{bmatrix} : z_i \in \mathbf{C},\ |z_i| = 1 \right\}$$

of unitary diagonal matrices is abelian and isomorphic to $\mathbf{T}^n$, the n-fold power of $\mathbf{T}$, the unit circle in $\mathbf{C}$. The unitary characters (irreducible representations) of A define a group isomorphic to $\mathbf{Z}^n$. They may be explicitly described by the formula

$$(3.5.4.2) \qquad \chi_m(a) = \prod_{n=1}^{n} z_i^{m_i},$$

where $m = (m_1, m_2, \dots, m_n)$ is an n-tuple of integers, and

$$a = a(z) = \operatorname{diag}(z_1, z_2, \dots, z_n)$$

is the diagonal matrix with diagonal entries $z_i \in \mathbf{T}$. The characters $\chi_m(a)$ form an orthonormal basis for $L^2(A)$ with respect to Haar measure on A

(assuming, as we will, that the Haar measure is normalized so that the total volume of A is 1).

Spectral theory for unitary matrices (cf. **[Lang3, Stra]**, etc.) tells us that every unitary matrix is conjugate to a diagonal matrix. Thus the map

$$\Gamma : A \times \mathrm{U}_n \to \mathrm{U}_n, \qquad \Gamma(a, g) = gag^{-1} \tag{3.5.4.3}$$

is surjective. It is clear that $\Gamma(a, gb) = \Gamma(a, g)$ for $b \in A$. Hence the map Γ actually factors to

$$\widetilde{\Gamma} : A \times (\mathrm{U}_n/A) \to \mathrm{U}_n.$$

The factored map $\widetilde{\Gamma}$ is generically finite-to-one. On the open dense set of matrices with n distinct eigenvalues, it is an $n!$-to-one covering map: two diagonal matrices define the same conjugacy class in U_n if and only if one can be turned into the other by permuting its diagonal entries. One can think of $\widetilde{\Gamma}$ as defining a system of "polar coordinates" on U_n.

Let dg denote Haar measure on U_n. Since $\overline{\gamma}$ is finite-to-one, up to sets of measure zero we can use it to lift dg up to $A \times (\mathrm{U}_n/A)$. Thus we can find a unique measure $d\mu(a, \dot{g})$ on $A \times (\mathrm{U}_n/A)$ such that the set where $\widetilde{\Gamma}$ is singular has measure zero and such that on the set where $\widetilde{\Gamma}$ is finite-to-one, we have the formula

$$\int_{A\times(\mathrm{U}_n/A)} f(a, \dot{g})\, d\mu(a, \dot{g}) = \int_{\mathrm{U}_n} \left(\sum_{x\in\widetilde{\Gamma}^{-1}(g)} f(x) \right) dg \tag{3.5.4.4}$$

for f a function on $A \times (\mathrm{U}_n/A)$.

The coset space U_n/A also possesses a left-invariant measure $d\dot{g}$. Since Haar measure on U_n is also conjugation invariant, we see $d\mu$ must be a product measure of the form

$$d\mu(a, \dot{g}) = d\nu(a) d\dot{g}.$$

Since we are in a context of smooth manifolds and smooth maps, we can easily believe that $d\nu$ is absolutely continuous with respect to Haar measure da on A:

$$d\nu(a) = \nu(a)\, da$$

for an appropriate function ν on A.

If we think of the map $\widetilde{\Gamma}$ as partitioning U_n into a family, parametrized (redundantly) by A, of fibers which are copies of U_n/A, then $\nu(a)$ tells us the volume of the fiber through a. This volume can be computed, up to a constant factor, as the determinant of an appropriate Jacobian mapping, which can be identified with the action $(1 - \mathrm{Ad}\, a)_{|\mathbf{a}^\perp}$ of a acting by conjugation on $\mathbf{a}^\perp$, the orthogonal complement to $\mathbf{a}$ in $\mathbf{u}_n$. (Note that, concretely, $\mathbf{a}^\perp$ is the space of skew-adjoint $n \times n$ matrices with zeros on the diagonal.) It is easy to compute that **[Wall2, Knap2, HoTa]**

$$\nu(a) = c|\det(1 - \mathrm{ad}\, a_{|\mathbf{a}^\perp})| = c \prod_{1\leq i<j\leq n} |z_i - z_j|^2$$

for an appropriate constant c. Here the z_i are the diagonal entries of a, as in formula (3.5.4.2). We will write

$$D(a) = \prod_{1 \le i < j \le n} (z_i - z_j). \tag{3.5.4.5}$$

(The function D is known as the discriminant; in our context it will play the role of the denominator in Weyl's character formula for U_n.) Then our formula for $\nu(a)$ can be written as

$$\nu(a) = cD(a)\overline{D(a)}. \tag{3.5.4.6}$$

Here $\overline{D(a)}$ denotes the complex conjugate of $D(a)$.

In formula (3.5.4.4), let us take the function f to be a pull-back from U_n by $\widetilde{\Gamma}$:

$$f(a, \dot{g}) = \phi(\widetilde{\Gamma}(a, \dot{g})) = \phi(\dot{g}a\dot{g}^{-1})$$

for some function ϕ on U_n. Taking into account the discussion above, we see

$$\#(W)\int_{\mathrm{U}_n} \phi(g)\,dg = \int_{A\times(\mathrm{U}_n/A)} \phi(\dot{g}a\dot{g}^{-1})\nu(a)\,da\,d\dot{g},$$

where $W \simeq S_n$ here indicates the group of permutations—the Weyl group of U_n. Suppose further that φ is invariant under conjugation. Then our formula simplifies to become

$$\int_{\mathrm{U}_n} \phi(g)\,dg = \frac{c}{\#(W)}\int_A \phi(a)D(a)\overline{D(a)}\,da. \tag{3.5.4.7}$$

This formula has a nice interpretation in terms of L^2-spaces. Let φ_1, φ_2 be two conjugation invariant functions. Setting $\varphi = \varphi_1\overline{\varphi}_2$ gives

$$\int_{\mathrm{U}_n} \varphi_1(g)\overline{\varphi_2(g)}\,dg = \frac{c}{\#(W)}\int_A (\varphi_1 D)(a)\overline{(\varphi_2 D)(a)}\,da. \tag{3.5.4.8}$$

Let $L^2(\mathrm{U}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ be the space of conjugation-invariant L^2-functions on U_n. The restriction of $\varphi \in L^2(\mathrm{U}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ to A will be invariant under the Weyl group $W = S_n$ of permutations of the diagonal coordinates. On the other hand, the discriminant function D is easily seen to be completely antisymmetric in the z_i; more precisely we have

$$D(p(a)) = \mathrm{sgn}(p)D(a), \qquad a \in A,\ p \in S_n, \tag{3.5.4.9}$$

where $\mathrm{sgn}\colon S_n \to \pm 1$ is the sign character: $\mathrm{sgn}(p)$ is 1 if p is an even permutation, and $\mathrm{sgn}(p)$ is -1 if p is odd. Thus the mapping

$$M_D : \varphi \to D\varphi \tag{3.5.4.10}$$

will take W-invariant or "symmetric" functions to "skew-symmetric" functions, i.e., functions transforming under W by the sign character. Let $L^2(A)^{W,\mathrm{sgn}}$ denote the subspace of skew-symmetric functions in $L^2(A)$. Let

$$\mathrm{res}_A : f \to f_{|A} \tag{3.5.4.11}$$

denote the restriction map from functions on U_n to functions on A. With this notation, we may express formula (3.5.4.8) as follows:

(3.5.4.12) The map $M_D \circ \mathrm{res}_A : L^2(U_n)^{\mathrm{Ad}\, U_n} \to L^2(A)^{W, \mathrm{sgn}}$ is, up to a scalar factor, a unitary isomorphism.

The discriminant function D is distinguished among all skew-symmetric functions by the property that it divides any one of them. More precisely, if f is a smooth skew-symmetric function on A, then we can write $f = D\varphi$ and the quotient φ, which will obviously be a symmetric function, will also be smooth. To see this, it suffices to consider two variables at a time: to show, say, that if f changes sign when z_1 and z_2 are interchanged, then f is divisible by $z_1 - z_2$. This can be done, for example, in terms of Fourier series. The basic formula is

$$z_1^n - z_2^n = (z_1 - z_2)(z_1^{n-1} + z_1^{n-2} z_2 + \cdots + z_2^{n-1}).$$

This argument shows that, in fact, if f has a finite Fourier series, then φ will also.

One way to create skew-symmetric functions is to take an arbitrary function f on A and skew-symmetrize it. Thus given f, we define

$$\text{(3.5.4.13)} \qquad \mathrm{skew}(f)(a) = \sum_{p \in W} \mathrm{sgn}(p) f(p(a)).$$

It is simple to check that $\mathrm{skew}(f)$ is skew-symmetric, and if f is already skew-symmetric, then $\mathrm{skew}(f) = \#(W) f$.

Consider $\mathrm{skew}(\chi_m)$ for some character χ_m of A, as in formula (3.5.4.2). The Weyl group also acts naturally on characters, by permuting the coordinates of the n-tuple m labeling χ_m. Specifically we have

$$w(m) = (m_{w^{-1}(1)}, m_{w^{-1}(2)}, \ldots, m_{w^{-1}(n)})$$

and

$$\chi_{w(m)}(w(a)) = \chi_a.$$

From these formulas, it is clear that

$$\text{(3.5.4.14)} \qquad \mathrm{skew}(\chi_{w(m)}) = \mathrm{sgn}(w)\, \mathrm{skew}(\chi_m).$$

Thus in constructing the functions $\mathrm{skew}(\chi_m)$, we need only consider m modulo the action of W. Thus let us define

$$\text{(3.5.4.15)} \qquad \widehat{A}^+ = \{\chi_m : m_1 \geq m_2 \geq \cdots \geq m_n\}.$$

It is easy to check that any character can be transformed by some element of W to a unique element of $\widehat{A}^+$. Thus we need only consider $\mathrm{skew}(\chi_m)$ for χ_m in $\widetilde{A}^+$.

We can also see from equation (3.5.4.14) that if any two coordinates of m are equal, then $\mathrm{skew}(\chi_m) = 0$. Thus in fact it is sufficient to consider $\mathrm{skew}(\chi_m)$ for χ_m belonging to

$$\text{(3.5.4.16)} \qquad \widehat{A}^{++} = \{\chi_m : m_1 > m_2 > \cdots > m_n\}.$$

Using elementary facts about Fourier series, we can see

(3.5.4.17) The functions $\#(W)^{-(1/2)}\operatorname{skew}(\chi_m)$, $m \in \widehat{A}^{++}$, define an orthonormal basis for $L^2(A)^{W, \mathrm{sgn}}$.

Let us remark that $\widehat{A}^{++}$ has a minimal element. That is, if we define

(3.5.4.18) $$\rho = (n-1, n-2, \ldots, 1, 0)$$

then

(3.5.4.19) $$\widehat{A}^{++} = \chi_\rho \widehat{A}^{+} = \{\chi_\rho \chi_m = \chi_{\rho+m} : \chi_m \in \widehat{A}^{+}\}.$$

Since the functions $\operatorname{skew}(\chi_m)$ span the skew-symmetric functions, we must be able to express D (cf. (3.5.4.5)) as a linear combination of the $\operatorname{skew}(\chi_m)$. In fact, by considering which characters could possibly occur in the expansion of the product defining D, we can conclude

(3.5.4.20) $$D = \operatorname{skew}(\chi_\rho)$$

with ρ as in (3.5.4.18). This identity, which is equivalent to the evaluation of the Vandermond determinant

$$\det \begin{vmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & z_n & z_n^2 & \cdots & z_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (z_j - z_i) = (-1)^{n(n-1)/2} D,$$

is one of the most fertile in mathematics. In **[Macd2]**, I. G. Macdonald discovered a class of identities attached to afine root systems that turned out to be analogs of (3.5.4.20) for affine Kac-Moody Lie algebras **[Kac1]**. The developments of this theme are still proceeding at a rapid pace (cf. **[KaPe, Lepo1, 2, LeMi, Macd3, Gust, Heck 1–3, Morr, HeOp, Opda1–3, Zeil]**, etc.).

Before continuing, let us note one consequence of the identity (3.5.4.20): it allows us to explicitly determine the constant c in formulas (3.5.4.7) and (3.5.4.8). Indeed, formula (3.5.4.20) tells us that D is the sum of $n!$ characters of A, with coefficients ± 1. Since characters are orthonormal in $L^2(A)$, we conclude

$$\int_A |D|^2(a)\, da = n! = \#(W).$$

Using this and $\varphi = 1$, the constant function, in (3.5.4.7) tells us that $c = 1$. Hence formula (3.5.4.8) reads simply

(3.5.4.21) $$\int_{\mathrm{U}_n} \varphi_1(g)\overline{\varphi_2(g)}\, dg = \frac{1}{\#(W)} \int_A (\varphi_1 D)(a)\overline{\varphi_2 D(a)}\, da,$$

for $\varphi_1, \varphi_2 \in L^2(\mathrm{U}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$.

Now turn to consideration of the irreducible characters of U_n. These are the functions

(3.5.4.22) $$\mathrm{ch}_\rho(g) = \operatorname{trace}\ \rho(g),$$

where $\rho : \mathrm{U}_n \to \mathrm{GL}(V)$ is an irreducible representation of U_n on a finite-dimensional vector space V. Evidently a similar definition can be made for any compact group. The significance of the characters for the representation theory of compact groups is summarized by the Peter-Weyl Theorem (cf. **[HeRo, Knap2, PeWe, Loom]**, etc.). To state it we need some notation.

Let G be a compact group. Let dg be Haar measure on G, normalized so that G has total mass equal to 1. Let $L^2(G)$ be the L^2-space of G with respect to dg, and let $L^2(G)^{\mathrm{Ad}\,G}$ denote the subspace of conjugation-invariant functions.

We can convolve elements of $L^2(G)$ (cf. §A.1.12). It is easy to check, for any locally compact group, that the convolution of two L^2-functions is continuous. Since we have G compact, continuous functions are L^2, so L^2 is an algebra under convolution. In fact,

$$\begin{aligned}\|f_1 * f_2\|_2 = \left\|\int_G f_1(g)(L_g f_2)\,dg\right\|_2 &\leq \int_G |f_1(g)|\,\|f_2\|_2\,dg \\ &\leq \|f_1\|_1\,\|f_2\|_2 \leq \|f_1\|_2\|f_2\|_2.\end{aligned}$$

Here, $\|f\|_p$ denotes the L_p-norm of a function on G. The last inequality follows since G has total mass 1. It is easy to check that $L^2(G)^{\mathrm{Ad}\,G}$ is the center of $L^2(G)$.

Let V be a finite-dimensional vector space, and $\rho : G \to \mathrm{GL}(V)$ a representation of G on V. We can define the character of ρ, ch_ρ, by formula (3.5.4.22).

Recall that $\widehat{G}$ denotes the set of irreducible unitary representations of G.

THEOREM 3.5.4.23 (Peter-Weyl). *Let G be a compact group.*

(a) *Every continuous irreducible representation σ of G is finite dimensional and unitary (i.e., given irreducible σ acting on V, there is a G-invariant hermitian inner product on V).*

(b) *Every irreducible representation of G can be realized as a subrepresentation of the left regular representation on $L^2(G)$.*

(c) *The irreducible characters ch_σ, $\sigma \in \widehat{G}$, constitute an orthonormal basis for $L^2(G)^{\mathrm{Ad}\,G}$.*

(d) *The functions $e_\sigma = (\dim\sigma)\mathrm{ch}_\sigma$, $\sigma \in \widehat{G}$, are idempotents for the convolution algebra structure on $L^2(G)$. They are precisely the minimal central idempotents in $L^2(G)$. Thus we have a decomposition*

$$\begin{aligned}L^2(G) \simeq \sum_{\sigma\in\widehat{G}} e_\sigma * L^2(G) &= \sum_{\sigma\in\widehat{G}} e_\sigma * L^2(G) * e_\sigma \\ &= \sum_{\sigma\in\widehat{G}} \sigma \otimes \sigma^*\end{aligned}$$

of $L^2(G)$ into mutually orthogonal, minimal, two-sided ideals. Each ideal $e_\sigma * L^2(G) * e_\sigma$ is isomorphic to a matrix algebra of rank $\dim\sigma$, *and as a*

$G \times G$ module under left and right translation, is isomorphic to $\sigma \otimes \sigma^*$, where σ^* indicates the contragredient of σ.

REMARKS. (i) Part (b) is proved by considering matrix coefficients (see §A.1.11). Given part (b), part (a) is an application of the theory of integral operators and the spectral theorem for compact selfadjoint operators **[Lang2, RiNa]** (or for selfadjoint algebras of compact operators). Parts (c) and (d) are analogs of the Schur orthogonality relations for finite groups, and are proved in essentially the same way, via Schur's Lemma **[HeRo, Knap2, Lang3]**.

(ii) To quote the Peter-Weyl Theorem in the derivation of the Weyl character formula is unhistorical as **[PeWe]** appeared several years after **[Weyl1]**. However, it is natural.

With these preparations, we are ready to state

(3.5.4.24) *Weyl character formula for* U_n: The characters of the irreducible representations of U_n are the functions

$$\frac{\mathrm{skew}(\chi_{m+\rho})}{\mathrm{skew}(\chi_\rho)} = \frac{\mathrm{skew}(\chi_{m+\rho})}{D}, \qquad m \in \widehat{A}^+.$$

Here $\mathrm{skew}(\chi_m)$ is defined in formula (3.5.4.13).

PROOF. Indeed, we know that if σ is a representation of U_n, then $\sigma_{|A}$ will be a direct sum of irreducible representations, i.e., characters, of A. Thus $(\mathrm{ch}_\sigma)_{|A} = \mathrm{ch}_{(\sigma_{|A})}$ will be a *positive integer* linear combination of elements of $\hat{A}$. Also, of course, ch_σ is conjugation invariant, so $\mathrm{ch}_{\sigma_{|A}}$ is symmetric. Thus the product $D(\mathrm{ch}_{\sigma_{|A}})$ will be an integer linear combination of characters of A, and will be skew-symmetric. It follows easily that $D(\mathrm{ch}_{\sigma_{|A}})$ is an *integer* linear combination of the functions $\mathrm{skew}(\chi_m)$, $m \in \widehat{A}^{++}$.

On the other hand, we know from Schur orthogonality, Theorem 3.5.4.23 (c), that the norm of ch_σ in $L^2(\mathrm{U}_n)$ is 1. It follows from (3.5.4.21) that the norm of $D(\mathrm{ch}_{\sigma_{|A}})$ in $L^2(A)$ is $\#(W)^{1/2}$. Combining this with the previous paragraph forces $D(\mathrm{ch}_{\sigma_{|A}})$ to be $\pm\,\mathrm{skew}(\chi_m)$ for a single χ_m. The sign can be checked by inspecting the coefficient of $\chi_{m-\rho}$ in $\mathrm{skew}(\chi_m)/\mathrm{skew}(\chi_\rho)$, and seeing it is positive (in fact, it is 1). The fact that all χ_m in $\widehat{A}^{++}$ are needed to express the characters follows from the completeness part of Theorem 3.5.4.23(c).

REMARKS. (a) To me, this proof is simply magical. If you attempt to analyze it, it dissolves into a few simple calculations and some general nonsense-airy nothing.

(b) One can recognize the same objects appearing here as in §3.5.3. The set $\widehat{A}^+$ is the collection of dominant integral weights, ρ is the half-sum of the positive roots, the alternating sum in (3.5.4.24) mirrors the Euler characteristic of the exact sequence (3.5.3.21), there is the same phenomenon of shifting by ρ, etc.

(c) A remarkable feature of this proof is that it simply identifies the irreducible characters. What the representations associated with these characters might be is, for the purposes of this argument, irrelevant and ignored. Of course, what the modules are was known well before this argument was given, from the highest weight theory described in §3.5.3. However, for noncompact groups the situation was reversed: Harish-Chandra **[HaCh19, 20]** gave a construction of discrete series characters using methods extending those explained in §3.5.6, well before these modules were constructed **[OkOz, Schm1-3, Hott, Part1, Wall4]**. Further, the early explicit constructions of discrete series modules all depended on knowing the character. It was not until **[AtSc, FlJe, Wall2]** that the existence of discrete series representations was established independently of character theory.

(d) Of course, the representation with character $\text{skew}(\chi_{m+\rho})/\text{skew}(\chi_\rho)$ is the representation with highest weight χ_m.

(e) By a l'Hopital's Rule argument as $a \in A$ approaches the identity, one obtains from the character formula a formula for $\dim \sigma$ (cf. **[Weyl2, Knap2, Jaco1]**, etc.).

(f) In the case of U_n, the character formula is due to I. Schur, who used rather different arguments **[Schu]**.

3.5.5. The highest weight theory (cf. Theorem 3.5.3.14) and the Weyl character formula (3.5.4.24) are the main constituents of our understanding of representations of compact Lie groups. Both were in place by the mid-1920s, well before the invention of the orbit method. However, both have been given interpretations consistent with the orbit picture. Even these interpretations, which date mainly from the 1950s, preceded the formulation of the orbit picture, and they provided guidance for the development of the representation theory of noncompact semisimple groups. In this section we will discuss the realization of representations by means of cohomology of line bundles over flag varieties. This is often called the Borel-Weil theory, but its full articulation is due to Borel-Weil **[Serr3]**, Bott **[Bott]**, and Kostant **[Kost4]**.

An essential aspect of the BWBK theory is the double interpretation of flag manifolds as homogeneous spaces, either for compact groups, or for their complexifications. We will describe this in general terms and illustrate it for the special unitary group SU_n.

We will discuss the BWBK theory for a connected, simply-connected, semisimple compact Lie group K. This is the essential case; allowing K to be disconnected, non-simply-connected, or to be nonsemisimple (i.e., to have a positive-dimensional center) has mainly nuisance value: it complicates the discussion without requiring any essential ideas. It is for this reason that we use SU_n rather than U_n for our example.

Let K be a connected, simply-connected, semisimple compact Lie group with Lie algebra $\mathbf{k}$, let $\mathbf{g} = \mathbf{k}_{\mathbf{C}} \simeq \mathbf{k} \otimes \mathbf{C}$ be the complexification of $\mathbf{k}$, and let G be the simply-connected Lie group with Lie algebra $\mathbf{g}$. Since $\mathbf{g}$

is complex, G likewise will carry a complex structure. We may think of K as the subgroup of G whose Lie algebra is $\mathbf{k} \subseteq \mathbf{g}$. If $K = \mathrm{SU}_n$, then $G = \mathrm{SL}_n(\mathbf{C})$. Let $T \subseteq K$ be a maximal abelian subgroup (a Cartan subgroup, also called a maximal torus since it is a product of circles). This notation is inconsistent with that of the preceding and following sections; but here we have another use for A. Let $\mathbf{t}$ be the Lie algebra of T, $\mathbf{a} = \mathbf{t}_{\mathbf{C}}$ its complexification. For $K = \mathrm{SU}_n$, we may take $\mathbf{a}$ to be the complex diagonal matrices of trace zero, and $\mathbf{t}$ the pure imaginary ones. The algebra $\mathbf{a}$ is a Cartan subalgebra of $\mathbf{g}$, and we have the root space decomposition of $\mathbf{g}$, as described in formula (2.8.6). We may make a choice of positive roots, and the corresponding Borel subalgebra $\mathbf{b}^+$ (cf. formulas (3.5.3.2)–(3.5.3.5)). Let $B \subseteq G$ be the connected subgroup whose Lie algebra is $\mathbf{b}^+$. Since $\mathbf{b}^+$ is its own normalizer in $\mathbf{g}$, B is necessarily closed. For $K = \mathrm{SU}_n$, we may take B to be the complex upper triangular matrices of determinant 1.

The Iwasawa decomposition **[Knap2, Wall2]** and §A.2.3.5 for G says that

$$G = KB, \qquad B \cap K = T. \tag{3.5.5.1a}$$

For $G = \mathrm{SL}_n(\mathbf{C})$, this amounts to the Gram-Schmidt orthonormalization procedure in $\mathbf{C}^n$. We also have the factorization

$$B = AN^+, \tag{3.5.5.1b}$$

where A and N^+ are the connected subgroups of G whose Lie algebras are $\mathbf{a}$ and $\mathbf{n}^+$ (cf. formula (3.5.3.5)). The group A is the complexification of T; it is called a Cartan subgroup or maximal torus of G. Every character χ of T extends in a unique way to a holomorphic character (i.e., a group homomorphism which is holomorphic with respect to the complex structures on A and $\mathbf{C}^\times$):

$$\chi : A \to \mathbf{C}^\times. \tag{3.5.5.2}$$

The complexification process described above also establishes, by a process of differentiation and analytic continuation just as discussed above for T, $\mathbf{t}$, $\mathbf{a}$, and A, bijections among the following sets:

$$\begin{aligned}
&\{\text{irreducible unitary representations of } K\} \\
&\quad \leftrightarrow \{\text{irreducible complex representations of } \mathbf{k}\} \\
&\quad \leftrightarrow \{\text{irreducible complex linear representations of } \mathbf{g}\} \\
&\quad \leftrightarrow \{\text{irreducible holomorphic representations of } G\}.
\end{aligned}$$

By a complex linear representation of $\mathbf{g}$ we mean a complex linear homomorphism $\sigma : \mathbf{g} \to \mathrm{End}(V)$ of $\mathbf{g}$ into the endomorphisms of some complex vector space V. Complex linearity of σ guarantees that σ is determined by its restriction to the real form $\mathbf{k}$ of $\mathbf{g}$. Similarly, a holomorphic representation $\sigma : G \to \mathrm{GL}(V)$ is a representation which is holomorphic as a mapping of complex manifolds. It is easy to check that the representation σ of G

is holomorphic if and only if the associated representation of $\mathbf{g}$ is complex linear.

Given a character χ of T, consider the induced representation $C_c^\infty(T\backslash K; \chi)$ (cf. §A.1.14). (Since K is compact, the subscript c in C_c^∞ is superfluous.) There is a geometric interpretation of $C^\infty(T\backslash K; \chi)$ in terms of line bundles **[FeDo, Huse, GrHa]**; we will review it. The quotient mapping

$$\begin{array}{c} K \\ \big\downarrow \\ T\backslash K \end{array} \tag{3.5.5.3}$$

can be thought of as a principal fiber bundle **[Huse]** with fiber T. Given a representation ρ of T on a space V, we can form the associated vector bundle $V \times_\rho K$. If $\rho = \chi$ is one-dimensional, then we simply have a line bundle. Comparison of the definition of $V \times_\rho K$ with the definition of induced representation shows that the functions in $C^\infty(T\backslash K; \rho)$ may be thought of as sections of the vector bundle $V \times_\rho K$.

The decomposition (3.5.5.1a) shows that

$$T\backslash K \simeq B\backslash G. \tag{3.5.5.4}$$

Since $B\backslash G$, being a quotient space of complex groups, is a complex manifold, we may use identification (3.5.5.4) to think of $T\backslash K$ as a complex manifold. In the case $G = \mathrm{SL}_n(\mathbf{C})$, it is the set of all "complete flags" in $\mathbf{C}^n$: sequences of nested spaces $\{D\} = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = \mathbf{C}^n$, with $\dim V_j = j$. In general, $T\backslash K$ is called the (complete) *flag variety* of G. The action by right translations of K on $T\backslash K$ extends holomorphically to an action of G. Further, given a character $\chi \in \widehat{T}$, we may extend χ holomorphically to A, then to a character of B trivial on N^+. Having done that, we may consider the induced representation $C^\infty(B\backslash G; \chi)$ (see §A.1.14). Decomposition (3.5.5.1a) then shows that by restricting elements of $C^\infty(B\backslash G; \chi)$ to K, we obtain an isomorphism

$$C^\infty(B\backslash G; \chi) \simeq C^\infty(T\backslash K; \chi). \tag{3.5.5.5}$$

On the other hand, the line bundle $\mathbf{C} \times_\chi G$ is a holomorphic line bundle over $B\backslash G$. Denote it by L_χ. In these circumstances, it is natural to look at the space $\Gamma(B\backslash G; L_\chi) = H^0(B\backslash G; L_\chi)$ of *holomorphic* sections of L_χ; and more generally, one can consider the (Dolbeault or, equivalently, sheaf) cohomology groups $H^p(B\backslash G; L_\chi)$ **[GrHa, Hart]**. Since G acts holomorphically on L_χ, the spaces $H^p(B\backslash G; L_\chi)$ will all be G-modules. Note that $H^0(B\backslash G; L_\chi)$ is a subspace of $C^\infty(B\backslash G; \chi)$—the kernel of the $\overline{\partial}$ operator; however, the higher cohomology groups are not subspaces of $C^\infty(B\backslash G; \chi)$. The BWBK theory describes the spaces $H^p(B\backslash G; L_\chi)$ as G-modules, and relates this to Lie algebra cohomology.

The group $\widehat{T}$ of characters of T is a lattice, isomorphic to $\mathbf{Z}^r$, $r = \dim T$. Given $\chi \in \widehat{T}$, let $D\chi$ be the derivative of χ at the identity. Then $D\chi \in \mathbf{t}^* \subseteq \mathbf{a}^*$, and the map $\chi \to D\chi$ identifies $\widehat{T}$ with a lattice inside $\mathbf{t}_{\mathbf{C}}^* = \mathbf{a}^*$. We will pass back and forth between χ and $D\chi$ without comment. Holomorphic extension to A further identifies $\widehat{T}$ to a group of quasicharacters (i.e., homomorphisms into $\mathbf{C}^x$) of A. The differentials of these quasicharacters are elements of $\mathbf{a}^*$, the same elements $D\chi$, $\chi \in \widehat{T}$, previously obtained. Denote by $\mathbf{a}_{\mathbf{R}}^*$ the real linear span of the lattice of $D\chi$. Then $\mathbf{a}_{\mathbf{R}}^*$ is a real form of $\mathbf{a}$, that is, we have the decomposition

$$\mathbf{a}^* = \mathbf{a}_{\mathbf{R}}^* \oplus i\mathbf{a}_{\mathbf{R}}^* \tag{3.5.5.6}$$

of $\mathbf{a}^*$ as a real vector space. In the corresponding decomposition

$$\mathbf{a} = \mathbf{a}_{\mathbf{R}} \oplus i\mathbf{a}_{\mathbf{R}} \tag{3.5.5.7}$$

we have $\mathbf{t} = i\mathbf{a}_{\mathbf{R}}$; that is, the elements of $\mathbf{a}_{\mathbf{R}}$ are purely imaginary on $\mathbf{t}$. For $K = \mathrm{SU}_n$, the Lie algebra $\mathbf{a}$ consists of complex diagonal matrices of trace zero, $\mathbf{a}_{\mathbf{R}}$ is the subspace of real diagonal matrices, and $\mathbf{t}$ is the space of purely imaginary diagonal matrices.

The Weyl group of A is the normalizer of A in G, modulo A:

$$W \simeq N(A)/A, \tag{3.5.5.8}$$

where $N(A)$ is the normalizer of A in G. We may also describe W as the normalizer of T in K, modulo T:

$$W \simeq N(T)/T. \tag{3.5.5.9}$$

The group W acts on A, hence on $\mathbf{a}$, by pullback via the exponential map—this action is via linear transformations. By duality, W acts on $\mathbf{a}^*$. Under these related actions, T, $\mathbf{t}$, $\mathbf{a}_{\mathbf{R}}$, $\mathbf{a}_{\mathbf{R}}^*$, and the lattice $\widehat{T} \subseteq \mathbf{a}_{\mathbf{R}}^*$ are all preserved by W. The action of W on $\mathbf{t}$ or $\mathbf{a}_{\mathbf{R}}^*$ is generated by reflections in hyperplanes—these reflections are the elements of W contained in the copies of SL_2 generated by root subspaces $\mathbf{g}_\alpha$, $\mathbf{g}_{-\alpha}$, as described in §2. Also as described there, the reflection hyperplanes divide $\mathbf{a}_{\mathbf{R}}$ and $\mathbf{a}_{\mathbf{R}}^*$ into convex cones, the Weyl chambers, which are permuted simply transitively by W. The positive Weyl chamber in $\mathbf{a}_{\mathbf{R}}$, relative to $\mathbf{b}^+$, or B is

$$\mathbf{a}_{\mathbf{R}}^+ = \{a \in \mathbf{a} : \alpha(a) \geq 0, \text{ all } \mathbf{g}_\alpha \in \mathbf{b}^+\}, \tag{3.5.5.10}$$

and the corresponding positive chamber in $\mathbf{a}_{\mathbf{R}}^*$ is

$$(\mathbf{a}_{\mathbf{R}}^*)^+ = \{\lambda \geq 0 \text{ on } \mathbf{a}_{\mathbf{R}}^+\}. \tag{3.5.5.11}$$

For $K = \mathrm{SU}_n$, the Weyl group is S_n, the symmetric group, which acts by permuting the diagonal entries of elements of $\mathbf{a}_{\mathbf{R}}$, which consists of traceless real diagonal matrices, and $\mathbf{a}_{\mathbf{R}}^+$ is the cone in $\mathbf{a}_{\mathbf{R}}$ consisting of matrices whose diagonal entries a_i decrease with $i : a_i \geq a_{i+1}$.

Write

$$\widehat{T}^{+} = \widehat{T} \cap (\mathbf{a}_{\mathbf{R}}^{*})^{+}. \tag{3.5.5.12}$$

Then $\widehat{T}^{+}$ consists of the dominant characters, or dominant integral weights—the highest weights of finite-dimensional representations, as described in §3.5.3.

The result of Borel and Weil describes the cohomology of line bundles defined by inverses of dominant characters [**Serr**, **Warn**, **Knap2**].

THEOREM 3.5.5.13 (Borel-Weil). (a) *Let $\chi^{-1} \in \widehat{T}^{+}$ be a dominant character. (We then say χ is antidominant.) Then the space $H^0(B\backslash G; L_\chi)$ of global holomorphic sections of the line bundle L_χ over $B\backslash G$ defined by χ is an irreducible G- (or K-) module, isomorphic to the dual of the representation whose highest weight is χ^{-1}.*

(b) *For $p > 0$, $H^p(B\backslash G; L_\chi) = 0$.*

REMARKS. (a) Theorem 3.5.5.13 is connected to the orbit method through a double interpretation of the complexification of the Lie algebra of $\mathbf{k}$: one as the (real) Lie algebra of the complexified group G, i.e., *real* right-invariant vector fields on G, and one as *complex* right-invariant vector fields on K. For functions which are holomorphic on G, these two interpretations coincide. Thus, for a holomorphic section f of L_χ, left invariance of f by N^+, as a function on G, can be interpreted in terms of $f_{|K}$ as a condition of being annihilated by the complex vector fields on K defined by $\mathbf{n}^+ \subseteq \mathbf{k}_{\mathbf{C}}$. Put another way, a function in $C^\infty(T\backslash K; \chi)$ will extend to a *holomorphic* N^+-left-invariant function on G if and only if it is annihilated by the vector fields from $\mathbf{n}^+ \subseteq \mathbf{k}_{\mathbf{C}}$, which may be seen to define a system of Cauchy-Riemann type equations. Thus, interpreted on K, the holomorphy condition becomes a condition of being an eigenfunction for the algebra $\mathbf{b}^+ = \mathbf{n}^+ + \mathbf{a} \subseteq \mathbf{k}_{\mathbf{C}}$. The algebra $\mathbf{b}^+$ is seen to be a complex polarization for the element $i\chi \in \mathbf{t}^* \subseteq \mathbf{k}^*$, and this use of complex polarizations is closely analogous to the way they are used in [**AuKo**] to produce representations of solvable groups. Although the Auslander-Kostant construction can be replaced by a construction involving only the real group, but inducing from representations of Heisenberg groups, not just from characters, there does not seem to be any escape from infinitesimal constructions involving complex polarizations in the case of semisimple groups.

(b) The presence of inverses and duals in this result makes it somewhat confusing. Perhaps the quickest way to verify the proper formulation is to consider the element of $H^0(B\backslash G; L_\chi)^*$ defined by the Dirac δ at the identity in G. If ρ denotes the action of G on $H^0(B\backslash G; L_\chi)$ by right translations, then

$$\rho^*(b)(\delta)(f) = \delta(\rho(b)^{-1}f)$$
$$= \rho(b)^{-1}(f)(1) = f(b^{-1}) = \chi^{-1}(b)f(1) = \chi^{-1}(b)\delta(f),$$
$$(b \in B,\, f \in H^0)$$

whence

$$\rho^*(b)(\delta) = \chi^{-1}(b)\delta.$$

In other words, δ is a highest weight vector for $H^0(B\backslash G;\, L_\chi)^*$, with weight χ^{-1}.

Part (a), the positive part of this result, is essentially a restatement of the highest weight theory. The main observation is that $H^0(B\backslash G;\, L_\chi)$ can contain at most one N^+-invariant function. (Indeed, all of $C^\infty(B\backslash G;\, \chi)$ contains only one N^+-invariant function.) This is because N^+ has a dense orbit on $B\backslash G$. This follows from the Bruhat decomposition (see §1.1 and **[HaCh1, Knap2, Wall2]**) which says $G = BWN^+$ so that, in particular, there are only finitely many N^+ orbits on $B\backslash G$, one of which is open and dense; but the fact that there is an open N^+-orbit in $B\backslash G$ is more elementary than the Bruhat decomposition.

The fact that $H^0(B\backslash G;\, L_\chi)$ consists of holomorphic functions means that the G-invariant subspaces of $H^0(B\backslash G;\, L_\chi)$ are the same as the K-invariant subspaces—in particular $H^0(B\backslash G;\, L_\chi)$ must be a direct sum of irreducible finite-dimensional G-modules. The theorem of the highest weight means that any G-irreducible subspace of $H^0(B\backslash G;\, L_\chi)$ must contain an N^+-invariant vector. Hence, there can be at most one subspace. On the other hand, if V_χ is the irreducible $\mathbf{g}$-module with highest weight $-D\chi$ and highest weight vector v_χ, then by exponentiating the action of $\mathbf{g}$ we obtain an action of G on V_{-D_χ}, and the matrix coefficients (cf. §A.1.11)

$$\varphi_{\lambda, v_\chi}(g^{-1}) = \lambda(g^{-1}(v_\chi)) \qquad \lambda \in V^*_{-D_\chi}$$

define a G-equivariant embedding of $V^*_{-D_\chi}$ into, hence an isomorphism with, $H^0(B\backslash G;\, L_\chi)$.

The complementary part (b) of Theorem 3.5.5.13 is a consequence of the Kodaira Vanishing Theorem **[GrHa, Hart]**.

Theorem 3.5.5.13 provides a "geometric" realization of the irreducible representations of K (or $\mathbf{g}$). However, it also raises an issue that it only partially resolves: although we can form the line bundles L_χ over $B\backslash G$ for all χ in $\widehat{T}$, Theorem 3.5.5.13 only describes the cohomology groups $H^p(B\backslash G;\, L_\chi)$ for χ antidominant. The highest weight theory guarantees that $H^0(B\backslash G;\, L_\chi) = \{0\}$ if χ is not antidominant, but it is silent about higher cohomology. The structure of the higher cohomology groups was clarified by Bott [**Bott**]. To state Bott's result (conjectured by Borel and Hirzebruch), we need to introduce the character

$$\delta(a) = \det(\mathrm{Ad}(a)_{|\mathfrak{n}^+}) = \prod_{\alpha>0} \chi_\alpha(g). \tag{3.5.5.14}$$

Here χ_α is the character of A whose derivative $D\chi_\alpha$ is equal to the positive root α. Thus δ is the character of A whose derivative at the identity is the sum of the positive roots. We observe that δ is a holomorphic square root of the modular function of B,

$$d(\mathrm{Ad}(a)n) = |\delta(a)|^2\, dn \tag{3.5.5.15}$$

if dn is Haar measure on N^+. It is a somewhat subtle point in the structure theory of compact groups that, under our assumption that K is simply connected, δ itself has a square root in $\widehat{T}$; we will denote this by $\delta^{1/2}$. It is not hard to check that δ and hence $\delta^{1/2}$ is dominant. For example, if $G = \mathrm{SL}_n(\mathbf{C})$, and we use the usual diagonal coordinates $\{a_i\}$ on A, then

$$\begin{aligned}\delta(a_1, \dots, a_n) &= \prod_{i<j}(a_i a_j^{-1}) = \prod_j a_j^{n+1-2j} \\ &= \left(\prod_j a_j\right)^{n+1} \left(\prod_j a_j^{-j}\right)^2.\end{aligned} \tag{3.5.5.16a)}$$

Thus for SL_n, we have

$$\delta^{1/2}(a_1, \dots, a_n) = \prod_j a_j^{-j} = \prod_j a_j^{n-j}. \tag{3.5.5.16b}$$

Note that $\delta^{1/2}$ is essentially identical with the χ_ρ used in §3.5.4 (cf. formula (3.5.4.18)).

THEOREM 3.5.5.17 (Bott). *Consider $\chi \in \widehat{T}$, and form the associated holomorphic line bundle L_χ over $B\backslash G$.*

(a) *If $\chi\delta^{-1/2}$ is singular (i.e., fixed by a nontrivial element of W), then $H^p(B\backslash G; L_\chi) = 0$.*

(b) *If $\chi\delta^{-1/2}$ is not singular, then there is a unique w in W such that $w(\chi^{-1}\delta^{1/2})$ is dominant, i.e., $w(\chi\delta^{-1/2})$ is antidominant. In this case, set*

$$\psi = w(\chi^{-1}\delta^{1/2})\delta^{-1/2}. \tag{3.5.5.18}$$

Let $l(w) = l$ be the length of w as an element of W (cf. [**Hilr, Bour**]). *Then $H^p(B\backslash G; L_\chi) = 0$ for $p \neq l$, and*

$$H^l(B\backslash G; L_\chi) \simeq (V_\psi)^*, \tag{3.5.5.19}$$

where V_ψ is the irreducible representation of G with highest weight ψ, with ψ given by (3.5.5.18).

Bott's proof of Theorem 3.5.5.17 used spectral sequences. However, Bott noted that, by some elementary yoga in sheaf cohomology, this theorem is equivalent to a statement about the Lie algebra cohomology of the Lie algebra $\mathbf{n}^+$ of N^+ with coefficients in a $\mathbf{g}$-module V_ψ, $\psi \in \widehat{T}^+$. Since the Cartan subgroup A acts on $\mathbf{n}^+$ by automorphisms, one sees from the standard construction **[BoWa, Jacol, Knapl]** of Lie algebra cohomology that each cohomology group $H^p(\mathbf{n}^+, V_\psi)$ naturally has the structure of an A-module. Kostant **[Kost4]** gave a direct explicit description of $H^p(\mathbf{n}^+, V_\psi)$ as an A-module, obtaining Bott's result as a corollary. To state Kostant's Theorem, we introduce the notation $\mathbf{C}_\chi$ for the one-dimensional irreducible representation of A whose associated character is χ.

THEOREM 3.5.5.20 (Kostant). *Let $\psi \in \widehat{T}^+$ be a dominant character of A, and let V_ψ be the associated finite-dimensional irreducible representation. Then there is an A-module isomorphism*

$$H^q(\mathbf{n}^+, V_\psi) \simeq \sum_{l(w)=q} \mathbf{C}_{\tilde{w}(\psi)}, \tag{3.5.5.21}$$

where $\tilde{w}(\psi) = w(\psi\delta^{1/2})\delta^{-1/2}$.

The essential, and originally the most difficult, part of the proof is to show that only the $\mathbf{C}_{\tilde{w}(\psi)}$ can appear in the $H^q(\mathbf{n}^+, V_\psi)$. (Aribaud **[Arib]** gave a simplified argument based on the Weyl character formula.) This is now understood to be an aspect of the Harish-Chandra homomorphism **[Hump, Knap2, Wall2]**, which also accounts for the "ρ-shifts" in the $\tilde{w}(\psi)$. This basic result gives a precise description of the center of the universal enveloping algebra of $\mathbf{k}$, or $\mathbf{g}$. The direct sum decomposition (cf. (3.5.3.4))

$$\mathbf{g} = \mathbf{n}^+ \oplus \mathbf{a} \oplus \mathbf{n}^+$$

of $\mathbf{g}$ leads via the Poincare-Birkhoff-Witt Theorem (cf. **[Hump, Jacol, Serr2]**, etc.) to the decompositions

$$\begin{aligned} \mathcal{U}(\mathbf{g}) &\simeq \mathcal{U}(\mathbf{n}^-) \otimes \mathcal{U}(\mathbf{a}) \otimes \mathcal{U}(\mathbf{n}^+) \\ &\simeq \mathcal{U}(\mathbf{n}^-) \otimes \mathcal{U}(\mathbf{a}) \oplus \mathcal{U}(\mathbf{g})\mathbf{n}^+. \end{aligned} \tag{3.5.5.22}$$

For $u \in \mathcal{U}(\mathbf{g})$, denote by $p(u)$ the component of u in $\mathcal{U}(\mathbf{n}^-) \otimes \mathcal{U}(\mathbf{a})$, the first summand of decomposition (3.5.5.22).

THEOREM 3.5.5.23 (Harish-Chandra homomorphism). (a) *If $u \in \mathcal{U}(\mathbf{g})^{\mathrm{Ad}\,A}$, the subalgebra of $\mathcal{U}(\mathbf{g})$ of elements invariant under* $\mathrm{Ad}\,A$, *then $p(u) \in \mathcal{U}(\mathbf{a})$.*

(b) *The mapping $u \to p(u)$ defines an algebra homomorphism from $\mathcal{U}(\mathbf{g})^{\mathrm{Ad}\,A}$ to $\mathcal{U}(\mathbf{a})$.*

(c) *If we make the identifications*

$$\mathcal{U}(\mathbf{a}) \simeq S(\mathbf{a}) \simeq \mathcal{P}(\mathbf{a}^*),$$

where $S(\mathbf{a})$ *denotes the symmetric algebra on* $\mathbf{a}$ *and* $\mathscr{P}(\mathbf{a}^*)$ *the algebra of polynomials on* $\mathbf{a}^*$, *then the map*

$$\tilde{p} : \mathscr{Z}\mathscr{U}(\mathbf{g}) \to \mathscr{P}(\mathbf{a}^*) \tag{3.5.5.24(i)}$$

defined by

$$\tilde{p}(u)(\lambda) = p(u)(\lambda - \rho), \qquad \lambda \in \mathbf{a}^*, \tag{ii}$$

where

$$2\rho = D_\delta, \tag{iii}$$

i.e., ρ *is* $\frac{1}{2}$ *of the sum of the positive roots, is an isomorphism*

$$\tilde{p} : \mathscr{Z}\mathscr{U}(\mathbf{g}) \simeq \mathscr{P}(\mathbf{a}^*)^W \tag{3.5.5.25}$$

from the center of $\mathscr{U}(\mathbf{g})$ *to the algebra of Weyl group invariant functions on* $\mathbf{a}^*$.

Parts (a) and (b) of this theorem are proved by easy computations, while part (c) may be seen using the Verma module approach to the highest weight theory, as described in §3.5.3.

In an illustration, and in some sense the crucial case, of Theorem 3.5.5.23, we recall the formula

$$\begin{aligned}\mathscr{C} &= h^2 + 2(e^+e^- + e^-e^+) = h^2 + 2h + 4e^-e^+ \\ &= (h+1)^2 - 1 + 4e^-e^+\end{aligned}$$

for the Casimir element in $\mathscr{U}(\mathbf{sl}_2)$ (cf. §3.5.1). Note that $1 = (\frac{1}{2})\alpha^+(h)$, where α^+ is the positive root in $\mathbf{sl}_2$, since $[h, e^+] = \alpha^+(h)e^+ = 2e^+$.

The Harish-Chandra homomorphism impinges on Theorem 3.5.5.20 as follows. Given a representation ρ of $\mathbf{g}$ on a vector space V, the images $\rho(z)$, $z \in \mathscr{Z}\mathscr{U}(\mathbf{g})$, are operators which commute with $\rho(x)$, $x \in \mathbf{g}$, and in particular with $\rho(\mathbf{n}^+)$. It follows from the standard construction [**Jaco1**, **BoWa**, **Knap1**] of Lie algebra cohomology that $\rho(\mathscr{Z}\mathscr{U}(\mathbf{g}))$ will induce operators on the cohomology groups $H^q(\mathbf{n}^+, V)$. Thus the $\mathbf{n}^+$ cohomology of a $\mathbf{g}$-module may be regarded as a joint $\mathscr{Z}\mathscr{U}(\mathbf{g})$ and $\mathscr{U}(\mathbf{a})$-module. (The $\mathscr{U}(\mathbf{a})$-module structure is of course obtained as the infinitesimal version of the action of A.) Denote this action by ρ^+.

THEOREM 3.5.5.26 (Casselman-Osborne [**CaOs**, **Knap1**, **Voga2**]). *The action* ρ^+ *of* $\mathscr{Z}\mathscr{U}(\mathbf{g})$ *on* $H^*(\mathbf{n}^+, V)$ *factors through the Harish-Chandra homomorphism*:

$$\rho^+(u) = \rho^+(\tilde{p}(u)) \tag{3.5.5.27}$$

with $\tilde{p}$ *as in formula* (3.5.5.25).

This result follows from the general machinery of cohomology, if one observes that the standard resolution of V as a $\mathbf{g}$-module [**Jaco1**, **BoWa**, **Knap1**] is also a resolution of V as an $\mathbf{n}^+$-module, since $\mathscr{U}(\mathbf{g})$ is free as a module over $\mathscr{U}(\mathbf{n}^+)$ by Poincaré-Birkhoff-Witt [**Hump**, **Jaco1**, **Serr2**]. Its relevance for Theorem 3.5.5.20 is that, if V is irreducible, then $\rho(\mathscr{Z}\mathscr{U}(\mathbf{g}))$ consists of scalars, and via ρ^+ will obviously act by the same scalars. Thus formula (3.5.5.27) constrains the action of $\mathscr{U}(\mathbf{a}^*)$. Indeed, it immediately implies that the only characters of A which could possibly appear in formula (3.5.5.21) are the ones which do. As mentioned above, this is the essential step in the proof of Theorem 3.5.5.20, which in turn implies the "geometric realization" Theorems 3.5.5.13 and 3.5.5.17.

3.5.6. In this subsection we complete the geometric quantization version of the basic representation theory of compact groups by giving Harish-Chandra's orbit method interpretation of the Weyl character formula [**HaCh3**]. With hindsight one can see in this remarkable paper the seeds of a large fraction of nonabelian harmonic analysis as it has developed in the ensuing 30 years. Besides Harish-Chandra's own work on the construction of the discrete series, it foreshadows the whole orbit method and also implicitly uses the oscillator representation [**Foll1**, **Howe3**, **Shal**, **Weil1**]. Our account will make this last connection explicit. (The first explicit use of the connection is [**Verg2**]).

As in §3.5.4, we will present only the example of the unitary group U_n, to save notation and preparation. The Lie algebra $\mathbf{u}_n$ of U_n is the space of skew-adjoint $n \times n$ complex matrices. To be definite we recall

$$\mathbf{u}_n = \{T \in M_n(\mathbf{C}) : T = [t_{jk}];\ t_{kj} = -\bar{t}_{jk}\}, \tag{3.5.6.1}$$

where the overbar denotes complex conjugation, and $\{t_{jk}\}$, $1 \le j, k \le n$, are the entries of the $n \times n$ matrix T. The unitary group U_n acts on $\mathbf{u}_n$ by conjugation. As usual, we denote this action by Ad:

$$\mathrm{Ad}\, g(T) = gTg^{-1}, \qquad T \in \mathbf{u}_n,\ g \in \mathrm{U}_n.$$

On $\mathbf{u}_n$ we can define a positive definite inner product $(\ ,\ )$ by the formula

$$\begin{aligned}(S, T) &= \mathit{trace}(ST^*) = -\operatorname{trace}(ST) \qquad (S, T \in \mathbf{u}_n)\\ &= -\sum_{1\le j,k\le n} s_{jk}\bar{t}_{jk}\\ &= -\sum s_{jj}t_{jj} + 2\sum_{1\le j<k\le n} (\operatorname{Re} s_{jk}\operatorname{Re} t_{jk} + \operatorname{Im} s_{jk}\operatorname{Im} t_{jk}).\end{aligned} \tag{3.5.6.2}$$

This inner product is easily seen to be invariant under $\mathrm{Ad}\,\mathrm{U}_n$. Using $(\ ,\ )$ we can define a Fourier transform on functions on $\mathbf{u}_n$ by one of the usual recipes

$$\hat{f}(S) = \int_{\mathbf{u}_n} f(T)e^{-2\pi i(S,T)}\, dT. \tag{3.5.6.3}$$

Here dT is Lebesgue measure defined by coordinates with respect to any orthonormal basis for $(\ ,\ )$. For example, we could take the coordinates it_{jj}, $2^{-1/2}\operatorname{Re} t_{jk}$, and $2^{-1/2}\operatorname{Im} t_{jk}$, $1 \le j < k \le n$. With this normalization of Lebesgue measure, the Fourier transform is unitary.

Let $\mathbf{a} \subseteq \mathbf{u}_n$ be the subspace of diagonal matrices

$$\mathbf{a} = \left\{ \begin{bmatrix} ib_{11} & & & \\ & ib_{22} & & \\ & & \ddots & \\ 0 & & & ib_{nn} \end{bmatrix} : b_{jj} \in \mathbf{R} \right\}. \tag{3.5.6.4}$$

We denote the typical element of $\mathbf{a}$ by B, and the typical entries of B will be ib_{jj}. The restriction of the inner product $(\ ,\ )$ of formula (3.5.6.2) defines an inner product on $\mathbf{a}$; in fact, it is just the standard Euclidean inner product with respect to the coordinates b_{jj}. The orthogonal complement $\mathbf{a}^{\perp}$ of $\mathbf{a}$ with respect to $(\ ,\ )$ is the space of skew-adjoint matrices with zeros on the diagonal. We can define the Fourier transform for functions on $\mathbf{a}$ by an analog of formula (3.5.6.3).

We know by spectral theory for self-adjoint matrices **[Lang3]** that every T in $\mathbf{u}_n$ is conjugate by U_n to an element of $\mathbf{a}$. Thus we have a surjective mapping

$$\begin{aligned} &\gamma : \mathbf{a} \times \mathrm{U}_n \to \mathbf{u}_n, \\ &\gamma(B, g) = gBg^{-1}, \qquad B \in \mathbf{a},\ g \in \mathrm{U}_n. \end{aligned} \tag{3.5.6.5}$$

This map is an infinitesimal analog of the map Γ of formula (3.5.4.3), and it has a basic theory parallel to the theory for Γ. First, it factors to a "polar coordinates" map

$$\tilde{\gamma} : \mathbf{a} \times (\mathrm{U}_n/A) \to \mathbf{u}_n$$

which is generically $n!$-to-one. Second, there is a polar-coordinates integration formula analogous to formula (3.5.4.21):

$$\int_{\mathrm{U}_n} \varphi_1(T)\overline{\varphi_2(T)}\, dT = c_o \int_{\mathbf{a}} \varphi_1 D(B)\overline{\varphi_2 D(B)}\, dB. \tag{3.5.6.6a}$$

Here c_o is an appropriate constant and D is, as before, the discriminant function

$$D(B) = \prod_{1 \le i < j \le n} (b_{ii} - b_{jj}), \qquad B \in \mathbf{a}. \tag{3.5.6.6b}$$

The constant c_o can be determined explicitly (see **[HaCh3, HoTa]**). As opposed to the situation in §3.5.4, here $D(B)$ is not to be thought of as a sum of characters, but as a polynomial function. Its structural interpretation is that it is the product of the positive roots for $\mathbf{a}$. The proof of formula (3.5.6.7) is parallel to that for (3.5.4.21). In particular the calculation of the volume $|D^2(B)|$ for the orbit $\operatorname{Ad}\mathrm{U}_n(B)$ is a Jacobian determinant computation slightly simpler than but quite similar to the volume factor $\nu(a)$. See the discussion preceding formula (3.5.4.5).

Also in parallel to §3.5.4, we may define the spaces $L^2(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ of conjugation invariant L^2 functions on $\mathbf{u}_n$, and $L^2(\mathbf{a})^{W,\mathrm{sgn}}$ of skew-symmetric functions on $\mathbf{a}$. We may define a map

$$M_D \circ \mathrm{res}_{\mathbf{a}} : L^2(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n} \to L^2(\mathbf{a})^{W,\mathrm{sgn}} \tag{3.5.6.7}$$

with notation parallel to statement (3.5.4.12), and it will again be true that this map (multiplied by $c_o^{1/2}$, with c_o as in (3.5.6.6a)) is a unitary isomorphism.

Since conjugation by U_n preserves the inner product $(\ ,\ )$, it will commute with the Fourier transform. Consequently, the space $L^2(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ will be invariant under the Fourier transform on $\mathbf{u}_n$. Similarly $L^2(\mathbf{a})^{W,\mathrm{sgn}}$ will be invariant under the Fourier transform on $\mathbf{a}$. Since we will now be considering the Fourier transform on $\mathbf{u}_n$ and on $\mathbf{a}$ at the same time, we will use the notations ${}^{\wedge_{\mathbf{u}}}$ and ${}^{\wedge_{\mathbf{a}}}$ respectively for them in order to be definite about which one is meant.

Harish-Chandra's discovery about the map $M_D \circ \mathrm{res}_{\mathbf{a}}$ was that it intertwines the two Fourier transforms.

THEOREM 3.5.6.8 (Harish-Chandra Restriction Theorem). *The mapping* (3.5.6.7) *satisfies*

$$M_D \circ \mathrm{res}_{\mathbf{a}} \circ {}^{\wedge_{\mathbf{u}}} = i^{-n(n-1)/2}\ {}^{\wedge_{\mathbf{a}}} \circ M_D \circ \mathrm{res}_{\mathbf{a}}.$$

In other words,

$$D(B)(\varphi^{\wedge_{\mathbf{u}}})(B) = i^{-n(n-1)/2}(D\varphi_{|\mathbf{a}})^{\wedge_{\mathbf{a}}}(B).$$

Since the Fourier transform is a nonlocal operator, a result like Theorem 3.5.6.8 is quite surprising. We will see shortly how special the circumstances are which give rise to this phenomenon.

To appreciate the structure underlying the Harish-Chandra Restriction Theorem, consider the Laplace operator on $\mathbf{u}$ dual to the inner product (3.5.6.2). It is the second-order, constant coefficient operator Δ, or $\Delta_{\mathbf{u}}$ when more specificity is needed, given by the formula

$$\Delta = \Delta_{\mathbf{u}} = \sum_{j=1}^{n} \frac{\partial^2}{\partial s_{jj}^2} + \frac{1}{2} \sum_{1 \le j < k \le n} \left(\frac{\partial^2}{\partial r_{jk}^2} + \frac{\partial^2}{\partial s_{jk}^2} \right), \tag{3.5.6.9}$$

where we take $t_{jk} = r_{jk} + i s_{jk}$, i.e., r_{jk} and s_{jk} are respectively the real and imaginary parts of t_{jk}. The factor $\frac{1}{2}$ occurs because, as noted above (see formulas (3.5.6.2) to (3.5.6.4)), the coordinates for which $(\ ,\ )$ looks like the standard Euclidean inner product are s_{jj}, $2^{-1/2}\, r_{jk}$ for $j < k$, and $2^{-1/2}\, s_{jk}$ for $j < k$. The operator $\Delta_{\mathbf{u}}$ is the standard Laplacian with respect to these coordinates.

Let us write

$$r^2(T) = (T, T). \tag{3.5.6.10}$$

Also let r^2 denote the operation of multiplication by r^2. It is easy to compute the commutator

$$[\Delta, r^2] = 4\left(\sum_{j=1}^{n} s_{jj}\frac{\partial}{\partial s_{jj}} + \sum_{1\leq j<k\leq n} r_{jk}\frac{\partial}{\partial r_{jk}} + s_{jk}\frac{\partial}{\partial s_{jk}}\right) + 2n^2 \tag{3.5.6.11}$$
$$= 4E + 2n^2,$$

where E is the standard Euler degree operator on $\mathbf{u}_n$, which multiplies polynomials of degree m by m.

For $\mathbf{R}^n$, consider the operators

$$e^+ = \pi i r^2, \qquad e^- = \frac{i\Delta}{4\pi}, \qquad h = E + \left(\frac{1}{2}\right)n^2. \tag{3.5.6.12}$$

Analogous operators may be defined for any space endowed with an inner product, as we have done above for $\mathbf{u}_n$, and the statements below will hold also for such spaces. Using formula (3.5.6.11) and some other simple calculations, we can check that $e^{\pm}$ and h form a standard basis for a copy of the Lie algebra $\mathbf{sl}_2$.

THEOREM 3.5.6.13 (Shale [**Shal**]). *There is a unique representation ω of $\widetilde{\mathrm{SL}_2}(\mathbf{R})$, the two-fold cover of $\mathrm{SL}_2(\mathbf{R})$, on $L^2(\mathbf{R}^n)$ such that the image of the associated representation of $\mathbf{sl}_2$ (see* §A.1.13) *is the span of the operators* (3.5.6.12).

REMARKS. (a) The operators (3.5.6.12) are a Lie subalgebra of the Lie algebra of all polynomial-coefficient differentials of total order (= polynomial degree + order of differentiation) two on $\mathbf{R}^n$. These operators are the span of

$$\pi i x_j x_k \quad \frac{1}{2}\left(x_j\frac{\partial}{\partial x_k} + \frac{\partial}{\partial x_k}x_j\right) = x_j\frac{\partial}{\partial x_k} + \frac{\delta_{jk}}{2} \quad \frac{i}{4\pi}\frac{\partial^2}{\partial x_j \partial x_k}. \tag{3.5.6.14}$$

This algebra is isomorphic to the symplectic Lie algebra in $2n$ variables, denoted $\mathbf{sp}_{2n}$. Shale actually showed there is a unitary representation of $\widetilde{\mathrm{Sp}}_{2n}(\mathbf{R})$, the two-fold cover of the real symplectic group in $2n$ variables, such that the image of the associated representation of the Lie algebra is the span of the operators (3.5.6.14).

(b) Shale's interest was quantum field theory. Shortly after Shale, Weil [**Weil1**], motivated by Segal's work on automorphic forms, independently showed the existence of this representation. Weil also established the existence of an analogous representation for $\mathrm{Sp}_{2n}(F)$, the symplectic group in $2n$ variables with values in a p-adic field F. Weil showed that this representation underlies the classical theory of θ-series, one of the most widely used means for constructing automorphic forms (cf. [**Igus**, **Shmz1**, **2**, **KuMi1–3**, **ToWa1–2**, **Shim**, **Shin2**, **Niwa**], etc.).

(c) I call this representation the *oscillator representation*, because of its close association with the quantum harmonic oscillator (see §3.1). Other names in use are the Weil representation, the Segal-Shale-Weil representation, the harmonic representation, etc.

(d) In some sense the oscillator representation is the quintessential example of geometric quantization, and it is derelict not to present its construction in detail. On the other hand, the construction gets rather involved and involves some special ideas, and so would constitute a sizeable digression. Also, I have written quite a bit about it **[Howe1–6]**, and do not wish to repeat myself here. Detailed accounts of it, from the viewpoint of geometric quantization, can be found for example in **[Blat, LiVe]**. My own account, which takes a somewhat different viewpoint, is **[Howe3]**.

(e) In fact, we do not need the full Theorem 3.5.6.13 for this discussion. We only need to exponentiate the operator $\frac{\Delta}{4\pi} - \pi r^2$, which is a multivariable variant of the Hamiltonian for the quantum harmonic oscillator discussed in §3.1. It can be handled by the same techniques. Thus our discussion is more or less complete on this point. However, Theorem 3.5.6.13 seems to identify the natural relevant structure for this situation. This connection was pointed out by Vergne in **[Verg2]**.

The relevance of Theorem 3.5.6.13 to the Harish-Chandra Restriction Theorem is that the Fourier transform is almost an element of $\omega(\widetilde{\mathrm{SL}_2}(\mathbf{R}))$. Consider the element

$$\mathbf{k} = e^+ - e^- = \frac{i}{2}\left(2\pi r^2 - \frac{\Delta}{2\pi}\right) \tag{3.5.6.15}$$

in our copy of $\mathbf{sl}_2$. An easy computation shows that $\mathbf{k}$ generates the standard maximal compact subgroup SO_2 inside SL_2. On the other hand, from calculations just like those of §3.1, we know the eigenvalues and eigenvectors of $\mathbf{k}$. From the standard formulas for the Fourier transform on $\mathbf{R}^n$, viz.,

$$(2\pi i x_j f)^\wedge = -\frac{\partial}{\partial x_j}(\hat{f}), \qquad \left(\frac{\partial}{\partial x_j} f\right)^\wedge = 2\pi i x_j \hat{f}, \qquad (e^{-\pi r^2})^\wedge = e^{-\pi r^2},$$

we can deduce that the eigenvectors for $\mathbf{k}$ are also eigenvectors for the Fourier transform, and further that the Fourier transform can be written as

$$\wedge = i^{-n/2} \exp(\pi \mathbf{k}/2). \tag{3.5.6.16}$$

See for example **[HoTa, Howe3]**.

REMARK. The scalar factor $i^{-n/2}$ in equation (3.5.6.16) comes from the fact that the smallest eigenvalue of $\mathbf{k}$ is $\frac{1}{2}$ rather than zero. This fact is interpreted in quantum mechanics as the Uncertainty Principle **[Shan]**, and in quantum electrodynamics as the zero-point energy, or energy of the vacuum **[Thir]**. It also reflects the fact that ω is a representation of $\widetilde{\mathrm{SL}_2}(\mathbf{R})$, and not of $\mathrm{SL}_2(\mathbf{R})$.

In view of formula (3.5.6.16), Theorem 3.5.6.8 follows from

THEOREM 3.5.6.17. *The mapping* $M_D \circ \operatorname{res}_{\mathbf{a}}$ *of formula* (3.5.6.7) *intertwines the restriction of the oscillator representations* $\omega_{\mathbf{u}}$ *and* $\omega_{\mathbf{a}}$ *of* $\widetilde{\mathrm{SL}_2}(\mathbf{R})$ *on* $L^2(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ *and* $L^2(\mathbf{a})^{W,\mathrm{sgn}}$ *respectively. It defines a unitary* (*up to multiples*) *equivalence of* $\widetilde{\mathrm{SL}_2}(\mathbf{R})$ *modules.*

REMARKS. We should note that the operators (3.5.6.12) are all invariant under conjugation by orthogonal transformations, and therefore the oscillator representation, whose existence is asserted by Theorem 3.5.6.13, will commute with orthogonal transformations. Since both spaces $L^2(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ and $L^2(\mathbf{a})^{W,\mathrm{sgn}}$ are defined by how their elements transform under certain orthogonal transformations, each is invariant under the relevant oscillator representation. Hence the assertion of Theorem 3.5.6.17 at least makes sense.

We will sketch a proof of this theorem.

Since in both representations the image of the operator $\mathbf{k}$ has discrete spectrum with finite-dimensional eigenspaces, as is revealed by the computations of §3.1, easy technical arguments reveal it is enough to show that Theorem 3.5.6.17 is true infinitesimally, i.e., that the map $M_D \circ \operatorname{res}_{\mathbf{a}}$ intertwines the operators (3.5.6.12) for $\mathbf{u}_n$ with their counterparts for $\mathbf{a}$. To do this for the operator e^+ is trivial: one needs only the facts that restriction is a homomorphism for pointwise multiplication, and that pointwise multiplication of complex-valued functions is commutative. To check it for $\mathbf{h}$ is also very simple: it uses the fact that $\mathbf{a}$ is invariant under scalar multiplication in $\mathbf{u}_n$, and that D is homogeneous of degree $\frac{1}{2}(\dim \mathbf{u}_n - \dim \mathbf{a})$.

Thus the crucial calculation is to show that the map $M_D \circ \operatorname{res}_{\mathbf{a}}$ intertwines the two Laplacians $\Delta_{\mathbf{u}}$ and $\Delta_{\mathbf{a}}$. We would like to perform this calculation in a moderately general context, to illustrate the issues involved. Related calculations are given in **[Helg3, Helg4]**. See also **[HoTa]**.

Consider $\mathbf{R}^n \subseteq \mathbf{R}^{n+m}$. Use coordinates $x_1, \ldots, x_n$ on $\mathbf{R}^n$, and let $y_1, y_2, \ldots, y_m$ be the remaining coordinates on $\mathbf{R}^{n+m}$. Imagine we are giving a "nonlinear orthogonal projection"

$$(3.5.6.18) \qquad \begin{aligned} &\Phi : \mathbf{R}^{m+n} \to \mathbf{R}^n, \\ &\Phi(x, y) = (\phi_1(x, y), \phi_2(x, y), \ldots, \phi_n(x, y)). \end{aligned}$$

Precisely, the points of $\mathbf{R}^n$ should be fixed by Φ, and the fibers $\Phi^{-1}(x)$, $x \in \mathbf{R}^n$, should intersect $\mathbf{R}^n$ orthogonally. In formulas, these conditions are

$$(3.5.6.19a) \qquad \Phi(x, 0) = x,$$

$$(3.5.6.19b) \qquad \frac{\partial \Phi}{\partial y_j}(x, 0) = 0.$$

For the calculations below, we need only that Φ be defined on some open set intersecting $\mathbf{R}^n$.

Our prototype for Φ is of course the map from $\mathbf{u}_n$ to $\mathbf{a}$ which takes T to an element in $\mathbf{a}$ conjugate to T. Globally, this map is not well defined, but in a neighborhood of any regular point of $\mathbf{a}$ it is well defined, and satisfies conditions (3.5.6.19).

We want to take a function f on $\mathbf{R}^n$, pull it back by Φ to a function on $\mathbf{R}^{n+m}$, apply the Laplacian on $\mathbf{R}^{n+m}$, then restrict the result to $\mathbf{R}^n$. Let Δ_x be the Laplacian in the x variables, and Δ_y the Laplacian in the y variables, so the full Laplacian on $\mathbf{R}^{n+m}$ is $\Delta_x + \Delta_y$. We compute, for $x \in \mathbf{R}^n$,

$$\begin{aligned}
\Delta(f \circ \Phi)(x, 0) &= (\Delta_x + \Delta_y)(f \circ \Phi)(x, 0) \\
&= \Delta_x(f)(x) + \left(\sum_{l,j} \frac{\partial}{\partial y_l}\left(\frac{\partial f}{\partial x_j}\frac{\partial \phi_j}{\partial y_l}\right)\right)(x, 0) \\
&= \Delta_x(f)(x) + \left(\sum_{ljk} \frac{\partial^2 f}{\partial x_k \partial x_j}\frac{\partial \phi_j}{\partial y_l}\frac{\partial \phi_k}{\partial y_l} + \frac{\partial f}{\partial x_j}\frac{\partial^2 \phi_j}{\partial y_l^2}\right)(x, 0) \qquad (3.5.6.20)\\
&= \Delta_x(f)(x) + \sum_j (\Delta_y \phi_j)(x, 0)\,\frac{\partial f}{\partial x_j}(x).
\end{aligned}$$

We want to compare this with the result of conjugating the Laplacian on $\mathbf{R}^n$ by a function. Thus we select a function ψ on $\mathbf{R}^n$ and we compute

(3.5.6.21)

$$\begin{aligned}
(\psi^{-1}\Delta_x \psi)f(x) &= \psi^{-1}(x)\Delta_x(\psi f)(x) \\
&= \psi^{-1}(\Delta_x(f)(x)\psi(x) + 2\psi^{-1}(x)\sum_j \frac{\partial \psi(x)}{\partial x_j}\frac{\partial f(x)}{\partial x_j} \\
&\quad + \psi^{-1}\, f(x)\, \Delta_x(\psi)(x) \\
&= \Delta_x(f)(x) + 2\sum_j \frac{1}{\psi}\,\frac{\partial \psi}{\partial x_j}\,\frac{\partial f}{\partial x_j} + f(x)\frac{\Delta_x(\psi)(x)}{\psi(x)}.
\end{aligned}$$

Comparing formulas (3.5.6.20) and (3.5.6.21), we see that if these two operations are going to be equal, the equations

$$\frac{2}{\psi}\,\frac{\partial \psi}{\partial x_j} = \Delta_y \phi_j(x, 0), \qquad (3.5.6.22a)$$

$$\Delta_x(\psi) = 0 \qquad (3.5.6.22b)$$

must hold. But equation (3.5.6.22a) already determines ψ up to a scalar multiple. It will only be by some lucky accident that we find the ψ so determined to be harmonic, i.e., that condition (3.5.6.22b) also holds. (In addition, the $\Delta_y \phi_j$ need to satisfy an integrability condition in order for (3.5.6.22a) to have a solution.)

Let us compute the ψ satisfying (3.5.6.22a) for the case of the eigenvalue projection of the unitary group. Here $\mathbf{R}^n = \mathbf{a}$, and the orthogonal space is $\mathbf{a}^\perp$, on which we may take coordinates $2^{-1/2} r_{jk}$ and $2^{-1/2} s_{jk}$, $1 \le j < k \le n$, where $t_{jk} = r_{jk} + i s_{jk}$ are the off-diagonal entries of a skew-adjoint matrix T. Let $\widetilde{E}_{jk}$ be the matrix with all entries zero except for ones in the (j, k)th and (k, j)th places. Let B be a diagonal matrix, with eigenvalues ib_l. To compute the right-hand side of formula (3.5.6.22a), we need to compute

$$\frac{d^2}{d\varepsilon^2} b_l(B + \varepsilon\widetilde{E}_{jk})_{|\varepsilon=0},$$

where here $ib_l(B+\widetilde{E}_{jk})$ indicates the lth eigenvalue of $B+\widetilde{E}_{jk}$, not the lth diagonal entry, which of course does not depend on ε.

It is easy to see that b_l does not change unless $l = j$ or $l = k$, and that the computation of b_j and b_k only involves the 2×2 matrix formed from the entries in the jth and kth rows and columns of $B + \varepsilon\widetilde{E}_{jk}$. Thus the computation comes down to a 2×2 matrix problem, viz, to find the eigenvalues of

$$\begin{bmatrix} \lambda_1 & \varepsilon \\ \varepsilon & \lambda_2 \end{bmatrix}.$$

We find that they are

$$\frac{1}{2}\left(\lambda_1 + \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + 4\varepsilon^2}\right)$$
$$\simeq \frac{1}{2}(\lambda_1 + \lambda_2 \pm |\lambda_1 - \lambda_2| \left(1 + \frac{2\varepsilon^2}{(\lambda_1 - \lambda_2)^2} + \cdots\right).$$

From this result, we easily find that

$$\Delta_y b_j = 2\sum_{k \neq j} \frac{1}{\lambda_j - \lambda_k}.$$

From this formula, we can see that $\psi = D$, the discriminant, will solve the system (3.5.6.22a). Further, it is well known that D, as the skew-symmetric polynomial of smallest possible degree, is necessarily harmonic. Thus D satisfies equations (3.5.6.22). Combined with our previous remarks, this establishes Theorem 3.5.6.17.

We now discuss the connection between Harish-Chandra's Restriction Theorem and the Weyl character formula. Consider the Schwartz spaces $\mathcal{S}(\mathbf{u}_n)$ and $\mathcal{S}(\mathbf{a})$ of rapidly decreasing smooth functions on $\mathbf{u}_n$ and $\mathbf{a}$. (See **[Foll, CoGr, Lang2]**, etc. for the basic facts on Schwartz spaces.) From Theorem 3.5.6.8 or Theorem 3.5.6.17 one can conclude that the unitary (up to scalars) map $M_d \circ \mathrm{res}_{\mathbf{a}}$ of formula (3.5.6.7) is also an isomorphism between the Schwartz spaces $\mathcal{S}(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}$ and $\mathcal{S}(\mathbf{a})^{W,\mathrm{sgn}}$. Dual to this map, we have a pullback map on tempered distributions:

$$(M_D \circ \mathrm{res}_{\mathbf{a}})^* : \mathcal{S}^*(\mathbf{a})^{W,\mathrm{sgn}} \to \mathcal{S}^*(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}. \tag{3.5.6.23}$$

Among the conjugation-invariant distributions on $\mathbf{u}_n$, probably the most important are the orbital integrals: given $T \in \mathbf{u}_n$, the *orbital integral* defined by the conjugacy class $\mathrm{Ad}\,\mathrm{U}_n(T)$ is

$$\mathcal{I}_T(f) = \int_{\mathrm{U}_n} f(\mathrm{Ad}\, g(T))\, dg. \tag{3.5.6.24}$$

Note that to get all possible orbital integrals, one need only consider T in $\mathbf{a}$. Every conjugation-invariant distribution on $\mathbf{u}_n$ is expressible as some sort of superposition of orbital integrals.

The analog of orbital integrals in the space $\mathscr{S}^*(\mathbf{a})^{W,\mathrm{sgn}}$ are the skew-symmetric sums

$$\text{skew}(\delta_B) = \sum_{w \in W} \mathrm{sgn}(w)\delta_{w(B)}, \qquad B \in \mathbf{a}, \tag{3.5.6.25}$$

where δ_B indicates the Dirac delta at B. Note that $\text{skew}(\delta_B) \neq 0$ if and only if $D(B) \neq 0$. From the elementary computation

$$\begin{aligned}(M_D \circ \mathrm{res}_{\mathbf{a}})^*(\text{skew}(\delta_B))(f) &= \text{skew}(\delta_B)(M_D \circ \mathrm{res}_{\mathbf{a}}(f)) \\ &= \text{skew } \delta_B(Df) = \#(W)D(B)f(B), \qquad f \in \mathscr{S}(\mathbf{u}_n)^{\mathrm{Ad}\,\mathrm{U}_n}, B \in \mathbf{a},\end{aligned}$$

we conclude

$$(M_D \circ \mathrm{res}_{\mathbf{a}})^*(\text{skew}(\delta_B)) = \#(W)D(B)\mathscr{I}_B. \tag{3.5.6.26}$$

We want to combine this formula with Theorem 3.5.6.8. Consider the Fourier transform of the orbital integral $\mathscr{I}_B$, $B \in \mathbf{a}$. Since $\mathscr{I}_B$ has compact support, its Fourier transform has the form

$$\widehat{\mathscr{I}_B} = (\widehat{\mathscr{I}_B})^0\, dT \tag{3.5.6.27}$$

where dT is Lebesgue measure on U_n and $(\widehat{\mathscr{I}_B})^0$ is an analytic function which is $\mathrm{Ad}\, U_n$-invariant, hence determined by its values on $\mathbf{a}$. Combining formulas (3.5.6.26) and (3.5.6.6) with Theorem 3.5.6.8, we conclude

$$(\widehat{\mathscr{I}_B})^0(B') = c_1(D(B)D(B'))^{-1} \text{ skew } \chi_{-B}(B') \qquad B' \in a \tag{3.5.6.28}$$

for an appropriate constant c_1. Here we have written $\chi_B(B') = e^{2\pi i(B,B')}$. (An extra computation shows that $c_1 = (\prod_{k=1}^{n-1} k!)(\frac{i}{2\pi})^m$ with $m = n(n-1)/2$.)

Using this formula, we can give an orbit-theoretic interpretation of the Weyl character formula. Let $\exp\colon T \to \exp(T)$ be the natural exponentiation map. The map exp allows us to identify a lattice in $\mathbf{a}$ with the character group of the torus $A = \exp \mathbf{a}$. Specifically, the restriction of exp to $\mathbf{a}$ is a group homomorphism. If $B \in \mathbf{a}$ is such that $\ker \chi_B \supseteq \ker \exp$, then χ_B may be pushed forward to A, where it will define a character. Let us call $B \in \mathbf{a}$ *integral* if χ_B factors through exp on $\mathbf{a}$. In terms of coordinates, we can see that if $B \in \mathbf{a}$ has diagonal entries ib_j, then B is integral if and only if the b_j's are integers. Further, if we identify B with its n-tuple of coordinates, then our notation χ_B for characters is consistent with the notation of §3.5.4.

The Weyl group W of permutations acts on $\mathbf{a}$ in the obvious way, and this action commutes with exp. We have the notion of positive Weyl chamber in $\mathbf{a}$ (cf. §2.10). In this case the positive Weyl chamber is

$$\mathbf{a}^+ = \{B \in \mathbf{a} : b_j \geq b_{j+1}\}.$$

Let us call B *dominant* if $B \in \mathbf{a}^+$. Denote by ρ the element of $\mathbf{a}$ whose jth diagonal entry is $i(n-j)$. (The parallel with formulas (3.5.4.18) and (3.5.5.16) will be evident. The need to multiply by i here comes simply from the concrete form of the Cartan subalgebra a.)

With these notations, we can state a formula which combines the Weyl character formula with the Harish-Chandra restriction formula.

THEOREM (3.5.6.29) (Harish-Chandra-Weyl character formula). *The irreducible characters of* U_n, *as functions on the maximal torus* $A = \exp \mathbf{a}$, *have the form*:

$$\mathrm{ch}_\sigma(\exp B') = \frac{D(B_{\sigma+\rho})D(B')(\widehat{\mathscr{I}_{B_{\sigma+\rho}}})^0(-B')}{c_1 D(\exp B')}, \tag{3.5.6.30}$$

where B_σ *is an appropriate dominant integral element of* $\mathbf{a}$.

REMARKS. (a) This formula is the analog for compact groups of the Kirillov character formula (3.3.1.7). A parallel for solvable groups which involves multiplication of the Fourier transform of the orbital integral by a correction factor is formula (3.4.1.2.1).

(b) Here again, as in the Verma module description of finite-dimensional representations (§5.3.3), and in the Weyl character formula (§5.3.4), we see a "ρ-shift" between the highest weight of the representation and the parameter we attach to the representation. Thus the $\mathrm{Ad}\,\mathrm{U}_n$-orbit associated to the trivial representation of U_n is not the origin, but rather the orbit through ρ. This phenomenon of ρ-shifts pervades the orbit method for semisimple groups. It is bookkeeping forced on us by the Harish-Chandra homomorphism (cf. Theorem 3.5.5.23).

(c) The argument given here for the Harish-Chandra Restriction Theorem and formula (3.5.6.30) looks quite different from the ones based on **[HaCh3]** (cf. **[Helg1, Wall2]**). Harish-Chandra first studies radial components of invariant differential operators, then uses them to deduce formula (3.5.6.30), then finally proves the Restriction Theorem. We established the Restriction Theorem first, then deduced formula (3.5.6.30). We could easily also deduce the results on radial components from the Restriction Theorem. However, although the order of main results is different, the crucial step in both developments is the computation of the radial component of the Laplacian (formula (3.5.6.20) and the discussion following it). In **[HaCh3]**, the oscillator representation appears only implicitly, in the use of taking commutators with the Laplacian to convert an invariant polynomial into the dual constant coefficient operator.

3.6. *Noncompact semisimple groups.* Noncompact semisimple groups have received the bulk of researchers' attention in representation theory since World War II, beginning with the papers of Wigner **[Wign]**, Bargmann **[Barg1]**, Gelfand-Naimark **[GeNa]**, and Harish-Chandra **[HaCh0]**. Until the late

1960s, Harish-Chandra was a fairly lonely pioneer, but since then the field has attracted a substantial number of workers. Fundamental progress has been made, including Harish-Chandra's Plancherel Formula **[HaCh22]**, and the classifications of Bernstein-Beilinson **[BeBe]**, Langlands **[Lgld4]**, and Vogan **[Voga4]** of the nonunitary irreducible representations. But many interesting and even basic problems, such as the determination of the unitary dual, remain to be solved (see, however, **[Voga5**, **Barb**, **Tadi]**), and much of the work already done sits undigested and unapplied.

In this account, we can only summarize some of the high points. We try to emphasize analogies with the easier classes of groups already discussed, and in particular we try to formulate results in terms of the orbit method. However, we emphasize that the structure of geometric quantization is for the most part imposed a posteriori, and played little role in the original arguments. Nevertheless, David Vogan currently is trying to create an understanding of the unitary dual more or less explicitly based on an appropriate version of the orbit method **[Voga6]**.

Due to the greater length and technical involvement of the arguments establishing results about noncompact semisimple groups, we must for the most part omit them, and be content with stating results. Two very useful books for learning a large portion of the theory in its current form are **[Knap2]** and **[Wall2]**. We also refer to **[Voga1]** for a nice overview.

3.6.1. PRINCIPAL SERIES. The main concrete objects of study in the representation theory of noncompact semisimple groups are the principal series. As with many things, the meaning of the term "principal series" can vary slightly with context. We begin by describing the most elementary case.

Let G be a semisimple Lie group, and let $P_0 \subseteq G$ be a minimal parabolic subgroup (cf. §A.2.4). We have a decomposition

$$P_0 = M_0 A_0 N_0, \tag{3.6.1.1}$$

where N_0 is the unipotent radical of P_0 (cf. §A.2.4), a connected, simply connected nilpotent group; A_0 is an abelian group, connected and simply connected (i.e., isomorphic to $\mathbf{R}^m$ for $m = \dim A_0$), and such that under the adjoint action, A_0 acts by diagonalizable matrices with positive real eigenvalues; and M_0 is compact. The group N_0 is normal in P_0, and M_0 and A_0 centralize each other. If $G = \mathrm{SL}_n(\mathbf{R})$, then N_0 consists of the unipotent upper triangular matrices, A_0 is the group of diagonal matrices with positive entries and determinant one, and M_0 is the group of diagonal matrices with entries ± 1 and determinant one.

Let ψ be a quasicharacter of A_0 (a homomorphism from A_0 to $\mathbf{C}^\times$), and σ an irreducible representation of M_0. Note that σ is finite dimensional. If V is the space of σ, define the representation $\sigma \otimes \psi$ of P_0 on V by

$$\sigma \otimes \psi(man)(v) = \psi(a)\sigma(m)(v), \qquad m \in M_0,\ a \in A_0,\ n \in N_0,\ v \in V. \tag{3.6.1.2}$$

Let δ_{P_0} be the modular function of P_0 (cf. formula (A.1.15.3)). Define the *principal series* representation associated to σ and ψ to be the induced representation (cf. §§A.1.14–16).

$$\text{(3.6.1.3)} \qquad \mathrm{P.S.}(\sigma, \psi) = \mathrm{ind}_{P_0}^{G} \ \sigma \otimes (\psi \delta_{P_0}^{-1/2}).$$

The set of representations $\mathrm{P.S.}(1, \psi)$, where 1 here denotes the trivial representation of M_0, is called the *spherical principal series.* (From an etymological viewpoint, this is a solecism: zonal principal series would be preferable.) The spherical principal series are slightly simpler than the $\mathrm{P.S.}(\sigma, \psi)$ with σ nontrivial, and they have some claim to a special place: they encompass all irreducible representations of G which contain a nonzero fixed vector for K, and consequently, they are the representations involved in the spectral analysis of functions on the symmetric space G/K. These topics are treated in detail in **[Helg1]** and **[GaVa]**.

The quasicharacters of A_0 form a complex vector space $\widehat{A}_0^{\mathbf{C}}$ of dimension $\dim A_0$. Thus if we let ψ vary in $\widehat{A}_0^{\mathbf{C}}$, the representations $\mathrm{P.S.}(\sigma, \psi)$ form a family, which in some sense (which can be made precise) is continuous or even holomorphic, of similar-looking representations. This is what the "series" in "principal series" connotes. Of course, $\delta_{P_0}^{-1/2}$ is a point in $\widehat{A}_0^{\mathbf{C}}$, and so multiplying ψ by $\delta_{P_0}^{-1/2}$ before forming the induced representation does not change the family of representations constructed, it only changes the way they are parametrized. The point of the chosen parametrization is that it takes unitary representations to unitary representations (cf. §§A.1.3 and A.1.16). For this reason multiplying by $\delta_{P_0}^{-1/2}$ before inducing is called *normalized induction* or *unitary induction.* The representations $\mathrm{P.S.}(\sigma, \psi)$ for ψ unitary are called the *unitary principal series*, and, by way of contrast, the whole principal series is sometimes called the *nonunitary principal series.* (We remark, however, that even for some nonunitary ψ, the representation $\mathrm{P.S.}(\sigma, \psi)$ can be given the structure of a unitary representation, though not in straightforward fashion [**Stei**, **Knap2**, p. 653].)

EXAMPLE. As an example, consider $G = \mathrm{SL}_2(\mathbf{R})$. We may take

$$\text{(3.6.1.4)} \qquad \begin{aligned} &P_0 = B = \left\{ \begin{bmatrix} a & x \\ 0 & a^{-1} \end{bmatrix} : a \in \mathbf{R}^{\times}, x \in \mathbf{R} \right\}, \\ &M_0 = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \qquad A_0 = \left\{ \begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} : a > 0 \right\}, \\ &N_0 = \left\{ \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} : x \in \mathbf{R} \right\}. \end{aligned}$$

Consider the space $\mathbf{C}^{\lambda, \varepsilon}(\mathbf{R}^2)$ of smooth functions on $\mathbf{R}^2 - \{0\}$ which are homogeneous of degree λ, $\lambda \in \mathbf{C}$, under positive dilations, and which are odd or even under reflection in the origin:

$$
\begin{aligned}
\mathbf{C}^{\lambda,\varepsilon}(\mathbf{R}^2) = \{f : (\mathbf{R}^2 - \{0\}) \to \mathbf{C},\ & f \text{ smooth},\\
& f(tx, ty) = t^{\lambda} f(x, y) \text{ for } t > 0, \text{ and}\\
& f(-x, -y) = \varepsilon f(x, y)\}
\end{aligned} \tag{3.6.1.5}
$$

for $\lambda \in \mathbf{C}$, $\varepsilon = \pm 1$. The action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathbf{R}^2$ gives rise in a natural way to an action ρ on $\mathbf{C}^{\infty}(\mathbf{R}^2 - \{0\})$ by the recipe

$$
\rho(g)(f)\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = f\left(g^{-1}\begin{bmatrix} x \\ y \end{bmatrix}\right). \tag{3.6.1.6}
$$

It is easy to check that the spaces $\mathbf{C}^{\lambda,\varepsilon}(\mathbf{R}^2)$ are invariant under ρ, so we may restrict ρ to any one of the $C^{\lambda,\varepsilon}$.

Define a mapping E from functions on $\mathbf{R}^2 - \{0\}$ to functions on $\mathrm{SL}_2(\mathbf{R})$ by the rule

$$
E(f)(g) = f\left(g^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = f\left(\begin{bmatrix} d \\ -c \end{bmatrix}\right), \quad f \in C^{\infty}(\mathbf{R}^2 - \{0\}),\ g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbf{R}). \tag{3.6.1.7}
$$

A straightforward calculation reveals that the mapping E defines an equivalence of $\mathrm{SL}_2(\mathbf{R})$ representations

$$
\mathbf{C}^{-\lambda-1,\varepsilon}(\mathbf{R}^2) \simeq \mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{\lambda}), \tag{3.6.1.8}
$$

where $\tilde{\varepsilon} : M_0 \to \{\pm 1\}$ is defined by $\tilde{\varepsilon}([\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}]) = \varepsilon$ and

$$
\tilde{a}^{\lambda}\left(\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}\right) = a^{\lambda}, \qquad a > 0.
$$

Thus the $\mathbf{C}^{\lambda,\varepsilon}(\mathbf{R}^2)$ serve as models for the principal series of $\mathrm{SL}_2(\mathbf{R})$.

Because of the homogeneity conditions (3.6.1.5) defining $\mathbf{C}^{\lambda,\varepsilon}(\mathbf{R}^2)$, we see that a function in this space is determined by its restriction to the unit circle

$$
S^1 = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 1\}
$$

and this restriction must be either an even or an odd function according as ε is $+1$ or -1. The circle S^1 is an orbit for the maximal compact subgroup $K = \mathrm{SO}_2$ of $\mathrm{SL}_2(\mathbf{R})$. Thus the Fourier series of $f_{|S^1}$ describes the decomposition of f into irreducible subspaces (in this example, eigenspaces, since SO_2 is abelian) for K. In particular, we see that each representation of K occurs with multiplicity at most one. For general groups, the Iwasawa decomposition (cf. equation A.2.3.5) shows us that the restriction to K of $\mathrm{P.S.}(\sigma, \psi)$ is also an induced representation

$$
\mathrm{P.S.}(\sigma, \psi)_{|K} \simeq \mathrm{ind}_{M_0}^{K}\, \sigma. \tag{3.6.1.9}
$$

By Frobenius reciprocity ([**HeRo**, **Knap2**, **Jaco2**] etc.), we may conclude that an irreducible representation τ of K occurs in P.S.(σ, ψ) with multiplicity equal to the multiplicity with which σ occurs in the restriction $\tau_{|M_0}$ of τ to M_0. This is certainly not more than $\dim\tau$. Thus all representations of K (= " K-types") occur in the principal series with finite multiplicity, which means that the principal series are admissible (cf. §3.6.5) representations [**Knap2**, **Wall2**].

The importance of the principal series is brought out by the following result.

THEOREM 3.6.1.10. (a) *The principal series representations* P.S.(σ, ψ) *all have finite composition series. The number of composition factors is bounded independently of* σ *and* ψ. *For fixed* σ, P.S.(σ, ψ) *is irreducible for a dense open set of* ψ.

(b) *Let* ρ *be any t.c.i.* (*cf.* §A.1.7) *representation of* G. *Then* ρ *is infinitesimally equivalent* (*cf.* §A.1.20) *to a subrepresentation of* P.S.(σ, ψ) *for appropriate* σ *and* ψ.

A weaker version of part (b), only asserting that σ could be realized as a constituent, i.e., subquotient, of some P.S.(σ, ψ), was proved by Harish-Chandra in early work [**HaCh4**], and later simplified by Lepowsky [**Lepo3**] and Rader (see also [**Wall2**]). The refinement giving σ as a subrepresentation was a long-standing problem, resolved by Casselman (see [**CaMi**]), using a refined version of Harish-Chandra's study [**HaCh13**] of the asymptotics of matrix coefficients. This study was based on the observation that elements of the center of the enveloping algebra give rise to differential equations which the matrix coefficients must satisfy. The differential equations imply that the matrix coefficients of an irreducible representation have certain asymptotic behavior at ∞ on G; this asymptotic behavior identifies the principal series into which the representation may be embedded.

The generic irreducibility of P.S.(σ, ψ), and finiteness of the composition series in general, has a fuzzier history. Generic irreducibility of the unitary principal series was proved by Bruhat [**Bruh**]. Finiteness of the composition series follows from Harish-Chandra's Regularity Theorem for characters [**HaCh14–18**] (see also [**Wall2**, **Vara**]), but proved this way, it is a deep result. Wallach [**Wall2**] gives a proof using his "Jacquet module." The composition series of $P(\sigma, \psi)$ when at least one constituent is finite dimensional is described by Vogan's extension of the Kazhdan-Lusztig formulas [**KaLu1**, **Voga7**]. Explicit examples and refinements have been given by Casian and Collingwood [**CaCo**, **Coll**]. However, there is still much to understand regarding the structure of these easily constructed representations.

Shortly after Casselman's proof of part (b), Langlands [**Lgld4**] (see also [**Knap2**, **Wall2**]) showed, again on the basis of Harish-Chandra's study of asymptotics of matrix coefficients, that by using a more general notion of principal series one can obtain a more or less canonical realization of a general irreducible representation. This is the "Langlands classification," which we will describe in §3.6.4. Here we describe the more general family of representations.

Let $P \subseteq G$ be any parabolic subgroup (cf. §A.2.4). Let $P = MAN$ be the Langlands decomposition ([**Knap2**, **Wall2**] and §A.2.4) of P, where N is the unipotent radical of P, A is a connected, simply connected abelian group, and M is semisimple. The group MA is the centralizer of A in G, and is a Levi component (cf. [**Jaco1**] and §A.2.4) for P. Let σ be an irreducible t.c.i. representation of M, and ψ a quasicharacter of A. Let δ_P be the modular function of P. We can define a representation $\sigma \otimes \psi$ of P in direct analogy with formula (3.6.1.2). Then we define the *generalized principal series* representation associated to σ and ψ to be

$$\text{P.S.}(\sigma, \psi) = \operatorname{ind}_P^G \sigma \otimes (\psi \delta_P^{-1/2}). \tag{3.6.1.11}$$

The parallel to formula (3.6.1.3) is patent, and sometimes one drops the adjective "generalized" and just calls the representations (3.6.1.11) "principal series." Note, however, that recipe (3.6.1.11) is much more of a black box than is (3.6.1.3), because the σ in (3.6.1.3) is a representation of the compact group M_0 and thus is to some extent understood, as described in §3.5. However, since M is noncompact, determination of the possible σ to stick in (3.6.1.11) is part of the problem under study, although for a smaller group. Further, we note that, by Harish-Chandra's Subquotient Theorem mentioned above, if σ is irreducible, the representations (3.6.1.11) are constituents of the usual principal series (3.6.1.3). For the Langlands classification, we only have to stick in for σ a special class of representations, the "tempered representations," to be described in §3.6.2.

We note again that (3.6.1.11) is a "normalized induction": if σ and ψ are unitary, then P.S.(σ, ψ) is also unitary.

We should also note that the formation of principal series, also known as *parabolic induction*, is eminently compatible with the orbit method. Let

$$\mathbf{g} = \mathbf{n}^- \oplus \mathbf{p} = \mathbf{n}^- \oplus \mathbf{m} \oplus \mathbf{a} \oplus \mathbf{n} \tag{3.6.1.12}$$

be the decomposition of the Lie algebra of G associated to the parabolic P. If the quasicharacter ψ in definition (3.6.1.6) is unitary, then it is the exponential of some $\lambda \in \mathbf{a}^*$ in the usual way:

$$\psi(\exp a) = e^{2\pi i \lambda(a)}, \qquad a \in \mathbf{a}.$$

(If ψ is not unitary, we could still use this formula, but would have to take λ in $\mathbf{a}_{\mathbf{C}}^*$, the complexified dual of $\mathbf{a}$.) Suppose we have associated the representation σ of M to the coadjoint orbit through some $\mu \in \mathbf{m}^*$, and to some polarization $\mathbf{q}$ of μ (which may well be a complex polarization, i.e., $\mathbf{q} \subseteq \mathbf{m}_{\mathbf{C}}^*$). Then it is easy to check that, for generic $\lambda \in \mathbf{a}^*$, the subalgebra $\mathbf{q} \oplus \mathbf{a}_{\mathbf{C}} \oplus \mathbf{n}_{\mathbf{C}}$ of $\mathbf{g}_{\mathbf{C}}$ will be a polarization of $\mu \oplus \lambda \in \mathbf{m}^* \oplus \mathbf{a}^* \subseteq \mathbf{g}^*$, and definition (3.6.1.11) would be the representation associated to the $\operatorname{Ad} G$-orbit of $\mu \oplus \lambda$ according to the usual orbit method yoga. In particular, if M_0 is discrete (which is the case for split real groups like $\mathrm{GL}_n(\mathbf{R})$, $\mathrm{Sp}_{2n}(\mathbf{R})$, etc.) or if M_0 is abelian (which is the case for complex groups, or quasisplit groups like $\mathrm{U}(n, n)$), then the unitary principal series of definition (3.6.1.3) are constructed via real polarizations. If M_0 is nonabelian, then we can understand its representations in terms of complex polarizations via the Borel-Weil Theorem (3.5.5.13), so the unitary principal series at least can be given a place in the orbit method. The combination of the Langlands classification (§3.6.4) and Zuckerman's derived functor construction (§3.6.5) extend this understanding to a class of representations at least big enough to write the Plancherel formula. A guide for further progress is provided by some conjectures of J. Arthur [**Arth2**], and the understanding of a family of representations dubbed unipotent [**BaVo**, **Vogal**], in homage to Lusztig's theory of representations of finite Chevalley groups [**Lusz1**]. Alternatively, the character theory of Harish-Chandra [**HaCh17–20**, **Wall2**, **Vara**], refined by Rossmann [**Ross1**, **2**] provides the direct connection of representations with orbits exemplified by formula (3.3.1.7).

3.6.2. TEMPERED REPRESENTATIONS. One of Harish-Chandra's basic insights into harmonic analysis on semisimple groups was the key role of what he called *tempered* representations. For his main goal, the Plancherel formula, the tempered representations were essential because, as he showed constructively, they are precisely the representations needed to perform the spectral analysis of $L^2(G)$. (This fact is now understood a priori [**CoHH**, **Bern3**]. It marks a fundamental difference between harmonic analysis on abelian, or even solvable groups, and semisimple groups.) They have turned out to be a basic ingredient in several other problems in representation theory, particularly problems suggested by automorphic forms [**Arth2**, **Sata**], and the Langlands classification (see §3.6.4).

Tempered representations are defined in terms of the decay of their matrix coefficients at ∞ on the group. The precise definition is in terms of a certain function Ξ introduced by Harish-Chandra [**HaCh10**] (see also [**Wall2**]). In fact Ξ is a natural function to consider: it is the matrix coefficient (cf. §A.1.11) associated to the (unique) K-invariant vector in P.S.$(1, 1)$, the spherical principal series associated to the trivial character of A_0. We can give an integral formula for Ξ as follows. Extend the modular function δ_{P_0} on P_0 to all of G by requiring the extended function to be invariant under left translation by K. Thus define

$$\delta_{P_0}(kp) = \delta_{P_0}(p), \qquad k \in K, p \in P_0. \tag{3.6.2.1}$$

Then Ξ is defined by

$$\Xi(g) = \int_K \delta_{P_0}^{-1/2}(gk)\, dk. \tag{3.6.2.2}$$

Here we take Haar measure on K to have total mass 1. The reader may wish to check that this is indeed the K-invariant matrix coefficient of P.S.$(1, 1)$.

Relevant definitions and formulas are given in §§A.1.11 and A.1.14–16. In any case, it is straightforward to check that Ξ is both left- and right-invariant by K, that $\Xi(g) \geq 0$, and that $\Xi(1) = 1$. Harish-Chandra established the following asymptotic properties of Ξ. We observe that by the Cartan decomposition (cf. formula (A.2.3.2)) the K-bi-invariance of Ξ implies that it is determined by its values on A_0^+, the positive Weyl chamber in A_0.

THEOREM 3.6.2.3. (a) *For $a \in A_0^+$, and any $\varepsilon > 0$, we have the estimates*

$$\delta_{P_0}^{-1/2}(a) \leq \Xi(a) \leq C_\varepsilon \delta_{P_0}^{-1/2+\varepsilon}(a)$$

for an appropriate constant C_ε.

(b) *For any $\varepsilon > 0$,*

$$\int_G \Xi^{2+\varepsilon}(g)\, dg < \infty.$$

Actually Harish-Chandra proves more refined estimates than these **[HaCh10, Knap2, Wall2, Vara]**; but these statements give the basic flavor of his results. Statement (b) is often phrased: "Ξ belongs to $L^{2+\varepsilon}(G)$."

Having Ξ in hand, we may define the notion of tempered representation. A representation ρ on the space V is called *tempered* if all its smooth matrix coefficients $\varphi_{u,\lambda}$ (cf. §A.1.11), $u \in V^\infty$, $\lambda \in (V^*)^\infty$, satisfy

$$\varphi_{u,\lambda}(g) \leq C_{u,v}\Xi(g) \tag{3.6.2.4}$$

for an appropriate constant depending on u and v. This is not precisely Harish-Chandra's definition, but is equivalent to it **[CoHH]**. It is also equivalent to requiring the smooth matrix coefficients to be $L^{2+\varepsilon}$.

3.6.3. DISCRETE SERIES. Just as tempered representations are essential to the spectral analysis of $L^2(G)$ for semisimple groups, the discrete series are essential to understanding tempered representations. Beyond this, discrete series are a fascinating phenomenon of general harmonic analysis. Also, discrete series for semisimple groups play a prominent role in the theory of automorphic forms **[BoWa, Lgld7, DeGW]**. The definition of discrete series makes sense for a general unimodular locally compact group G. Let ρ be an irreducible unitary representation of G on a Hilbert space $\mathscr{H}$. We call ρ a *discrete series* (or *square integrable*) if there exist u, v in $\mathscr{H}$ such that the matrix coefficient $\varphi_{u,v}$ (cf. §A.1.11) belongs to $L^2(G)$.

This rather innocent sounding definition has striking implications, described in the next result. The proof is a pleasant exercise in functional analysis, originally done by Godement **[Godel]** (see also **[Knap2]**).

THEOREM 3.6.3.1. *Let ρ be a discrete series representation of the unimodular locally compact group G. Let ρ be realized on the Hilbert space $\mathscr{H}$. Then the following assertions are true:*

(i) *Every matrix coefficient $\varphi_{u,v}$, $u, v \in \mathscr{H}$, of ρ is in $L^2(G)$.*

(ii) *There is a constant* d_ρ *such that*

$$\int_G \varphi_{u,v}(g)\overline{\varphi_{w,z}(g)}\, dg = \frac{(u, w)(z, v)}{d_\rho} \tag{3.6.3.2}$$

for any $u, v, w, z \in \mathscr{H}$. *Here* $(\ ,\ )$ *indicates the inner product in* $\mathscr{H}$.

(iii) *In particular, for fixed* $v \in \mathscr{H}$, *the mapping*

$$u \to \left(\frac{d_\rho}{(v, v)}\right)^{1/2} \varphi_{u,v}$$

defines an isometric G-intertwining from $\mathscr{H}$ *to a subspace of* $L^2(G)$. *Thus* ρ *is equivalent to a summand of* $L^2(G)$.

REMARKS. (a) Equation (3.6.3.2) is a generalization of the Schur orthogonality relations for finite or compact groups (cf. **[HeRo, Knap2, Jaco2]**, etc.). For those groups, if Haar measure is normalized to have total mass 1, the constant d_ρ is just the dimension of $\mathscr{H}$, also known of old as the degree of ρ. In the general case, when $\mathscr{H}$ is infinite dimensional, d_ρ is called the *formal degree* of ρ.

(b) The equation (3.6.3.2) is reminiscent of a fixed point formula: if $u = v = w = z$, then it expresses the integral of $|\varphi_{u,u}|^2$ as a multiple of $(u, u)^2 = |\varphi_{u,u}|^2(1)$, where 1 denotes the identity element of G. Indeed, formula (3.6.3.2) has a natural interpretation as a "trace formula."

With this background on discrete series for general groups, let us explain for semisimple Lie groups the relation between discrete series and tempered representations.

THEOREM 3.6.3.3. (a) *If* G *is a semisimple Lie group, and* ρ *is a discrete series representation of* G, *then* ρ *is tempered.*

(b) *Let* $P \subseteq G$ *be a parabolic subgroup with decomposition* $P = MAN$ *as in equation* (3.6.1.5). *If* σ *is a tempered representation of* M, *and* $\psi \in \widehat{A}$ *is a unitary character, then the principal series representation* P.S.(σ, ψ) (*cf. formula* (3.6.1.11)) *is tempered.* (*In brief: unitary parabolic induction preserves temperedness.*)

(c) *Every irreducible tempered representation of* G *is a summand of some* P.S.(σ, ψ) *as in part* (b), *where* σ *is a discrete series.*

The complete classification of tempered representations, i.e., a description of the precise decomposition of the tempered P.S.(σ, ψ), and the equivalences between the pieces, was given by Knapp and Zuckerman **[KnZu]**. Their results were given a nice orbit method interpretation by Rossmann **[Ross2]**, as an adjunct to his character formula, to be discussed in Theorem 3.6.3.7. However, it was known already from results of Bruhat in the 1950s **[Bruh]** for the basic principal series that for any σ and generic (i.e., for an open dense set of) ψ, the representation P.S.(σ, ψ) is irreducible. (It was to prove this that Bruhat studied the double coset decomposition now named after him.) These results were extended by Harish-Chandra to cover the case when P is nonminimal and σ is a discrete series. Hence Theorem 3.6.3.3 gives a description of "almost all" the tempered representations. In particular, the representations described by Theorem 3.6.3.3 are precisely the representations which enter into Harish-Chandra's Plancherel formula for G.

In sum, Theorem 3.6.3.3 shows that an understanding of tempered representations can to a large extent be reduced to an understanding of the discrete series.

Harish-Chandra **[HaCh19–20]** gave a description of the discrete series of a semisimple Lie group. He did so by explicitly constructing their characters, which he expressed in terms of Fourier transforms of invariant measures on certain orbits in the dual of the Lie algebra of G, a procedure with obvious analogies to the orbit method sketched above for nilpotent and solvable groups, and with Harish-Chandra's own formula for the characters of compact groups. That the parallel is essentially perfect, so that the characters of discrete series, and in fact of all tempered representations, can be described by a close cousin of the formula of Theorem 3.5.6.29, was established by W. Rossmann **[Ross1, 2]**. We will describe this in Theorem 3.6.3.7.

Since Harish-Chandra's classification of discrete series did not actually produce representations, in the sense of providing some concretely described spaces with some concretely given G-actions on them, a clamor soon arose for a "geometrical realization" of the discrete series. A candidate for such a realization, using "L^2-cohomology" and bearing strong analogies to the Bott-Borel-Weil Theorem (cf. §3.5.5) for compact groups was proposed by Langlands and Kostant **[Lgld7, Kost7]**. A realization of this sort was established in stages by W. Schmid **[Schm1–3]**. Other authors used variations on this theme to produce similar models for the discrete series (cf. **[Hott, OkOz, Part1, Wall4]**, etc.). However, all of these constructions depended on Harish-Chandra's existence proof via character theory; they did not independently establish either existence or exhaustion of the discrete series.

In the 1970s a more algebraic approach to problems of representation theory arose, and several purely algebraic constructions of discrete series and analogous representations were given **[EnVa, Part2, Zuck]**. The construction given by G. Zuckerman (see **[Voga, Wall2, Knap1]**) has turned out to be in some sense the most natural and has been shown to have numerous pleasant technical properties. It is now more or less the standard construction, and has been developed to the point where it can be used to give an independent proof of the existence of the discrete series **[Wall2]**. Proof that the construction exhausts all discrete series, however, still requires character theory. We describe Zuckerman's construction in §3.6.5.

Several other, rather different, ways to construct discrete series have also been developed. A construction of discrete series on the family of "semisimple symmetric spaces"—homogeneous spaces of the form G/H, where G is semisimple and H is the identity component of an automorphism of order 2 of G (see **[FlJe2, Berg]**)—was discovered by M. Flensted-Jensen **[FlJe2, Knap2]**. This proceeds by giving integral formulas for certain matrix coefficients of the representations, by means of a beautiful duality between pairs of semisimple symmetric spaces **[FlJe2]**. Since a semisimple group G can be thought of as the semisimple symmetric space $G \times G/\Delta(G)$, where

$$\Delta(G) = \{(g, g) : g \in G\}$$

is the diagonal in $G \times G$, Flensted-Jensen's construction yields discrete series for G as a special case. Flensted-Jensen's methods have been extended and to some extent completed by Oshima **[Oshi1, 2]**, using the theory of hyperfunctions and holonomic systems.

Another, again quite different, construction of the discrete series using an index theorem for covering spaces was given by M. Atiyah and W. Schmid **[AtSc]**.

Let us turn now to a concrete description of discrete series. First, we should note that not all semisimple groups have discrete series. A major insight of Harish-Chandra was that discrete series should be associated to compact Cartan subgroups. A Cartan subgroup of G is a subgroup whose complexified Lie algebra is a Cartan subalgebra of $\mathfrak{g}_{\mathbf{C}}$.

THEOREM 3.6.3.4. *Let G be a semisimple Lie group and $K \subseteq G$ a maximal compact subgroup. Then H has discrete series if and only if* rank $K =$ rank G, *i.e., iff a Cartan subgroup of K is also a Cartan subgroup in G iff G has a compact Cartan subgroup.*

Thus, for example, the rank of $\mathrm{SL}_n(\mathbf{R})$ is $n-1$, and that of its maximal compact subgroup SO_n is $[n/2]$. These are equal only for $n = 2$, so for $n \geq 3$, $\mathrm{SL}_n(\mathbf{R})$ has no discrete series. Since $\mathrm{SL}_n(\mathbf{R})$ occurs as a factor of the Levi component of maximal parabolics of many groups (e.g. of $\mathrm{O}_{p,q}$, $\mathrm{Sp}_{2n}(\mathbf{R})$), this very substantially cuts down on the number of parabolic subgroups one must worry about in the context of Theorem 3.6.3.3. Also, complex groups always have their compact real form as a maximal compact subgroup, and this always has rank equal to $\frac{1}{2}$ the rank of the full group (considered as a real Lie group), so complex groups have no discrete series. As a result, the only tempered representations for complex groups are constituents of the standard principal series induced from characters of the minimal parabolic subgroup. With hindsight we may say that it was this circumstance that permitted the early determination by Gelfand-Naimark **[GeNa]** and Harish-Chandra **[HaCh5]** of the Plancherel formula for complex semisimple groups.

Other examples are: $\mathrm{Sp}_{2n}(\mathbf{R})$ has maximal compact U_n, and both have rank n, so $\mathrm{Sp}_{2n}(\mathbf{R})$ has discrete series; $\mathrm{O}_{p,q}$ has rank $[\frac{p+q}{2}]$, while its maximal compact $\mathrm{O}_p \times \mathrm{O}_q$ has rank $[\frac{p}{2}] + [\frac{q}{2}]$, so $\mathrm{O}_{p,q}$ has discrete series if and only if at least one of p and q is even. In this connection, we may note that if p is odd, then $\mathrm{O}_{p,1}$, like complex groups, has only the standard principal series as tempered representations, and hence has a simple Plancherel formula, analogous to that for complex groups [**Wall3**].

Following Theorem 3.6.3.4, consider a semisimple Lie group G containing a compact Cartan subgroup T. The discrete series of G are then parametrized, more or less, by the characters of T, in a fashion similar to the description given in §§3.5.3–3.5.5 for the case of G compact. However, the "more or less" hides several tricky points, of which we will try to give some idea.

We will describe the discrete series by associating to them coadjoint orbits. This gives a formula for their characters, which is the original description given by Harish-Chandra [**HaCh19, 20**]. The refinement to the orbital description is due to Rossmann [**Ross1**]. For convenience, we will restrict our discussion to connected groups. The results can be extended to more general G, but at the expense of substantial technical fussing.

Let G be a semisimple group, connected and without center. Let $K \subseteq G$ be a maximal compact subgroup of G and $T \subseteq K$ a maximal torus (Cartan subgroup). We assume T is also a Cartan subgroup of G; this means (since G is connected) that T is its own centralizer in G. Let $\mathbf{g}$, $\mathbf{k}$, and $\mathbf{t}$ be the Lie algebras of G, K, and T respectively, and let $\mathbf{g}^*$, $\mathbf{k}^*$, and $\mathbf{t}^*$ be the duals of $\mathbf{g}$, $\mathbf{k}$, and $\mathbf{t}$. Via the Killing form (cf. equation (2.8.8)) on $\mathbf{g}$, we can identify $\mathbf{g}$ with $\mathbf{g}^*$, $\mathbf{k}$ with $\mathbf{k}^*$, and $\mathbf{t}$ with $\mathbf{t}^*$. We will not take full advantage of this identification, however, but merely use it to consider $\mathbf{t}^*$, which naturally is a quotient of $\mathbf{g}^*$, as a subspace of $\mathbf{g}^*$.

Let $N(T)$ be the normalizer of T in G. In fact, $N(T) \subseteq K$. Under standard technical assumptions [**GaVa; Vara**, p. 192)] which always hold if G is connected, the action of $N(T)$ on $\mathbf{t}$ (or $\mathbf{t}^*$) factors through the Weyl group W of $\mathbf{t}$ in K. We recall from §§2.9 and 2.10 that W is generated by reflection in certain hyperplanes $H_\alpha \subseteq \mathbf{t}$. These are the hyperplanes orthogonal to the roots α of $\mathbf{k}$ relative to $\mathbf{t}$; we will refer to them as *K-root* hyperplanes or *compact root* hyperplanes. The complement of the compact root hyperplanes is called the set of *K-regular elements.* Denote this set by $\mathbf{t}_{r,K}$. The connected components of $\mathbf{t}_{r,K}$ are open convex cones. These are permuted simply transitively by W. The closure of any one of them is called a *K-Weyl chamber.*

By duality, W acts on $\mathbf{t}^*$ also, and we use similar terminology to describe this dual action. The hyperplanes in $\mathbf{t}^*$ dual to the H_α will be denoted by H_α^*.

Since $\mathbf{t}$ is a Cartan subalgebra of $\mathbf{g}$, its complexification $\mathbf{t}_{\mathbf{C}}$ is likewise a Cartan subalgebra in $\mathbf{g}_{\mathbf{C}}$, the complexification of $\mathbf{g}$. In this context too, we have a Weyl group, which by rather flagrant abuse of notation we will indicate by $W_{\mathbf{C}}$. It is not hard to show that $W_{\mathbf{C}}$ preserves $\mathbf{t}$, considered as a real subspace of $\mathbf{t}_{\mathbf{C}} = \mathbf{t} \oplus i\mathbf{t}$. This is because $\mathbf{t}$ is characterized as the real subspace of $\mathbf{t}_{\mathbf{C}}$ on which the roots of $\mathbf{t}_{\mathbf{C}}$ in $\mathbf{g}_{\mathbf{C}}$ take on pure imaginary values. As a group of linear transformations of $\mathbf{t}$, the group $W_{\mathbf{C}}$ contains

W, and is also generated by reflections in hyperplanes. These hyperplanes will still be denoted by H_α, where now α is a root of $\mathbf{t}_\mathbf{C}$ in $\mathbf{g}_\mathbf{C}$. If H_α is a reflecting hyperplane of $W_\mathbf{C}$, but not of W, we call H_α a *noncompact root* hyperplane. The complement of all the H_α, compact or noncompact, is the set of G*-regular elements*, or just regular elements, denoted $\mathbf{t}_{r,G}$. We define a notion of Weyl chamber for $W_\mathbf{C}$ as for W. These are called *G-Weyl chambers*. Evidently each K-Weyl chamber contains several G-Weyl chambers. (To be precise, $\#(W_\mathbf{C}/W)$ of them.)

Again, we note we can dualize the above discussion to $\mathbf{t}^*$.

We can consider elements of $\mathbf{t}^*$ as defining unitary characters on $\mathbf{t}$ in the usual way: if $\lambda \in \mathbf{t}^*$, the associated character χ_λ of $\mathbf{t}$ is given by

$$\chi_\lambda(t) = e^{2\pi i\lambda(t)}, \qquad t \in \mathbf{t}.$$

Consider the exponential map $\exp: \mathbf{t} \to T$. Since T is commutative, exp is a group homomorphism. We call $\lambda \in \mathbf{t}^*$ *integral*, or T-integral if we need to specify T, if χ_λ factors through exp to define a character of T. As in §3.5.5, the set of integral λ form a lattice in $\mathbf{t}^*$, identified via the map $\lambda \to \chi_\lambda \circ \exp^{-1}$ to the Pontrjagin dual $\widehat{T}$ of T. Hence we denote it by $\widehat{T}$.

Let Σ be the set of roots of $\mathbf{t}_\mathbf{C}$ in $\mathbf{g}_\mathbf{C}$. Elements of Σ are the linear forms defined by eigenvectors for $\mathbf{t}_\mathbf{C}$ acting on $\mathbf{g}_\mathbf{C}$ via ad. As such they define by exponentiation characters of T, which we can then identify with elements of $\widehat{T} \subseteq \mathbf{t}^*$. As we have mentioned, the roots are elements in $\mathbf{t}_\mathbf{C}^*$ which most naturally take imaginary values on $\mathbf{t}$. To get them to be elements of $\mathbf{t}^*$, we have essentially multiplied them by i. (Note the i in the definition of χ_λ. This is a different convention from §3.5.5, where we did not multiply by i, instead we considered that $\widehat{T} \subseteq i\mathbf{t} \subseteq \mathbf{t}_\mathbf{C}$.)

Let $\mathscr{C} \subseteq \mathbf{t}$ be a G-Weyl chamber, and let $\Sigma_\mathscr{C}^+ \subseteq \Sigma$ denote the set of roots which take positive values on $\mathscr{C}$. We set

$$\rho_\mathscr{C} = \frac{1}{2} \sum_{\alpha \in \Sigma_\mathscr{C}^+} \alpha. \tag{3.6.3.5}$$

Although $\rho_\mathscr{C}$ clearly depends on the choice of $\mathscr{C}$, the difference $\rho_{\mathscr{C}_1} - \rho_{\mathscr{C}_2}$ for different chambers $\mathscr{C}_1$ and $\mathscr{C}_2$ is, by standard results [**Bour**, **Serr1**], a sum of elements of Σ. Hence the coset $\widehat{T} + \rho_\mathscr{C}$ of $\widehat{T}$ in $\mathbf{t}^*$ is independent of the choice of $\mathscr{C}$. We denote it by $\widehat{T} + \rho$.

The final ingredient we need before giving a precise description of the discrete series is an open set $U \subseteq \mathbf{g}$, which is connected, contains the origin, and is such that the exponential map $\exp: U \to G$ is a diffeomorphism onto its image. There is a natural maximal choice for U [**Wall2**] but we will not try to describe it. Let J be the Jacobian relating Haar measure on G with the push-forward via exp of Lebesgue measure on $\mathbf{g}$:

$$\int_{\exp U} f(\exp x)\, d(\exp x) = \int_U f(x) J(x)\, dx, \qquad f \in C_c^\infty(\exp U). \tag{3.6.3.6}$$

Here the measures $dg = d(\exp x)$ and dx are the appropriate Haar measures. Clearly J is smooth, conjugation-invariant, and positive, so it has a well-defined positive square root $J^{1/2}$, which also is smooth and conjugation-invariant.

THEOREM 3.6.3.7 (Harish-Chandra **[HaCh20]**, Rossmann **[Ross1]**). (a) *Let π be a discrete series representation of the connected semisimple Lie group G. There is a coadjoint orbit $\mathscr{O}_\pi \subseteq \mathbf{g}^*$ such that the character Θ_π of π may be computed from the Fourier transform of the invariant measure on $\mathscr{O}_\pi$ by means of the following formula:*

$$\Theta_\pi(f) = \text{trace } \pi(f) = \int_{\mathscr{O}_\pi} \left(\int_U f(\exp x) J^{1/2}(x) \overline{\chi_\lambda(x)}\, dx \right) d\lambda.$$

Here U is the neighborhood of 0 in $\mathbf{g}$ selected just above, $f \in C_c^\infty(\exp U)$, and $d\lambda$ is the appropriately normalized invariant measure on $\mathscr{O}_\pi$.

(b) *The orbit $\mathscr{O}_\pi$ intersects $\mathbf{t}^*$. The intersection $\mathscr{O}_\pi \cap \mathbf{t}^*$ consists of a W-orbit of points in $\widehat{T} + \rho$, and is contained in $\mathbf{t}^*_{r,G}$. The mapping*

$$\pi \to \mathscr{O}_\pi \to \mathscr{O}_\pi \cap \mathbf{t}^*$$

establishes a bijection between the discrete series of G and the W-orbits in $(\widehat{T} + \rho) \cap \mathbf{t}^*_{r,G}$.

REMARKS. (a) The parallel with the compact case (cf. §3.5.6) should be clear.

(b) Harish-Chandra **[HaCh14–18, Vara, Wall2]** showed that the character Θ_π of π (in fact of any irreducible representation of G) was given by integration against a locally L^1, conjugation-invariant function, analytic on the regular set. Furthermore, he gave an explicit formula for this function on T. We will describe his formula. Fix an element $\lambda \in \mathscr{O}_\pi \cap \mathbf{t}^*$, and let $\mathscr{C}$ be the G-Weyl chamber containing λ. Define the "denominator function"

$$D = \prod_{\alpha \in \Sigma^+_{\mathscr{C}}} (\chi_{\alpha/2} - \chi_{-\alpha/2}) = \sum_{s \in W_{\mathbf{C}}} \operatorname{sgn}(s) \chi_{s(\rho_{\mathbf{C}})}. \tag{3.6.3.8}$$

Here sgn is the standard sign character of $W_{\mathbf{C}}$. Although D depends on $\mathscr{C}$, the dependence is weak: the D associated to a different chamber equals this D up to a $\pm$ sign.

The character Θ_π is expressed, as a function on T, by the formula

$$\Theta_\pi = \pm \frac{\sum_{s \in W} \operatorname{sgn}(s) \chi_{s(\lambda)}}{D}. \tag{3.6.3.9}$$

Some words about interpreting this formula may be helpful. Because λ is in $\widehat{T} + \rho$, which may not equal $\widehat{T}$, it may happen that neither the numerator nor the denominator defines a function on T (i.e., factors through $\exp\colon \mathbf{t} \to T$). However, the quotient may also be written as

$$\Theta_\pi = \pm \frac{\sum_{s \in W} \operatorname{sgn}(s) \chi_{s(\mu+\rho)-\rho}}{\sum_{s \in W_{\mathbf{C}}} \operatorname{sgn}(s) \chi_{s(\rho)-\rho}}, \qquad \mu + \rho = \lambda,$$

and both the numerator and denominator of this expression do factor to T. In particular, Θ_π factors to T.

(c) The analogy between formula (3.6.3.9) and the Weyl character formula (cf. (3.5.4.24)) is clear. In fact, although the derivation of formula (3.6.3.9) is very substantially more difficult than that of (3.5.4.24), several key features of the argument are parallel. However, it should be noted that the denominator function D involves antisymmetrization over the "complex Weyl group" $W_{\mathbf{C}}$, and so has zeros along all the H_α, $\alpha \in \Sigma$, whereas the numerator of formula (3.6.3.9) involves only a sum over the "real Weyl group" W, and cannot be expected to vanish on the noncompact root hyperplanes. Thus the character Θ_π will have singularities.

(d) As in the case of compact groups (cf. §3.5.6), the numerator in formula (3.6.3.9) is provided by $\widehat{\mathscr{O}_\lambda}$, the Fourier transform of the orbital integral, and the denominator D is provided by the Jacobian factor $J^{1/2}$.

(e) The explicit formula (3.6.3.9) is due to Harish-Chandra **[HaCh19, 20]**. In the course of his argument, he established a less precise version of the orbital integral formula for $\widetilde{\Theta}_\pi$ given in Theorem 3.6.3.7: the expression for Θ_π was allowed to be a linear combination of the orbital integrals coming from coadjoint orbits through the full $W_{\mathbf{C}}$-orbit $W_{\mathbf{C}}(\lambda)$, rather than a single orbit. The more refined result, that only one orbit is involved in Θ_π, was established by Rossmann **[Ross1]**. The key step in Rossmann's analysis was an analog for noncompact G of Theorem 3.5.6.8.

There is a story that Harish-Chandra had considered whether a single-orbit expression like that of Theorem 3.6.3.7 might be valid, but was led to abandon such a hope by erroneous computations for the example of $\mathrm{SL}_2(\mathbf{R})$. If this is true, it gives a rare instance where Harish-Chandra's intuition, which guided him so well through the deep forest of semisimple harmonic analysis, led him astray. Particularly for $\mathrm{SL}_2(\mathbf{R})$, the distinctive analytic features of the various types of discrete series representations are so well mirrored by the geometry of the different type of elliptic coadjoint orbits, that the one-representation/one-orbit hypothesis seems, at least with hindsight, very plausible.

(f) The correct normalization of the invariant measure $d\lambda$ on $\mathscr{O}_\pi$ is given by the same universal normalization as in the nilpotent case, defined intrinsically in terms of the symplectic structure on $\mathscr{O}_\lambda$, as Kirillov **[Kiri]** suggested should hold.

(g) Rossmann **[Ross2]** has also shown that an orbital integral formula like that of Theorem 3.6.3.7 is valid for all tempered representations. In this generality, the bijectivity of the correspondence between representations and orbits breaks down. Sometimes different representations have characters which agree near the identity in G, and so correspond to the same orbit, and sometimes several orbits are needed to give the character of one representation. (This latter situation is exceptional, however.) These phenomena can already be seen in $\mathrm{SL}_2(\mathbf{R})$ **[Ross2]**.

(h) Duflo and Vergne **[DuVe1]** have showed that the orbital picture yields a nice interpretation of the Plancherel formula, yielding a sort of "Poisson-Plancherel" formula for G, partaking of the nature of both the classical Poisson and Plancherel formulas.

(i) The character formula (3.6.3.9) suggests what the restriction to K of the discrete series representation π should look like. We have observed that the numerator of (3.6.3.9) looks like the numerator of the Weyl character formula for K, while the denominator is the Weyl denominator for G, which contains the Weyl denominator for K as a factor. Decompose

$$\Sigma^+_{\mathscr{C}} = \Sigma^+_c \cup \Sigma^+_n, \tag{3.6.3.10}$$

where the Σ^+_c are the roots of $\mathbf{t}_{\mathbf{C}}$ acting on $\mathbf{k}_{\mathbf{C}}$ (the compact roots), and Σ^+_n are the remaining roots (the noncompact roots). Then

$$D_K = \prod_{\alpha\in\Sigma^+_c} (\chi_{\alpha/2} - \chi_{-\alpha/2}) \tag{3.6.3.11}$$

is the Weyl denominator for K, and

$$D_n = \prod_{\alpha\in\Sigma^+_n} (\chi_{\alpha/2} - \chi_{-\alpha/2}) = \pm\left(\prod_{\alpha\in\Sigma^+_n} \chi_{-\alpha/2}\right)\left(\prod_{\alpha\in\Sigma^+_n} (1-\chi_\alpha)\right) \tag{3.6.3.12}$$

is the quotient D/D_K. In analogy with the geometric series

$$\frac{1}{1-r} = \sum_{m=0}^{\infty} r^m$$

we may formally expand

$$\left(\prod_{\alpha\in\Sigma^+_n} (1-\chi_\alpha)\right)^{-1} = \sum p_n(\gamma)\chi_\gamma, \tag{3.6.3.13}$$

where γ is any weight of the form $\sum_{\alpha\in\Sigma^+_n} n_\alpha \alpha$, with nonnegative integers n_α, and $p_n(\gamma)$, the "partition function" for Σ^+_n, is the number of ways of expressing γ as such a sum. Note also that decomposition (3.6.3.10) induces a parallel decomposition (cf. formula (3.6.3.5))

$$\rho_{\mathscr{C}} = \rho_c + \rho_n. \tag{3.6.3.14}$$

With this notation, we can write $\prod_{\alpha\in\Sigma^+_n} \chi_{-\alpha/2} = \chi_{-\rho_n}$. It is easy to see that D_n is invariant under W, the Weyl group of K. Thus, for any $s \in W$, we may write formally

$$D_n = \chi_{-s\rho_n} \left(\sum p_n(\gamma)\, \chi_{s(\gamma)}\right)^{-1}.$$

Plugging this in to expression (3.6.3.9) gives us

$$(D_K)^{-1} \sum_{s\in W} \operatorname{sgn} s \sum_{\gamma} p_n(\gamma)\chi_{s(\lambda+\rho_n+\gamma)}. \tag{3.6.3.15}$$

If we count the number of times that a given K-dominant weight μ occurs in this sum, we obtain

$$\left(\sum_{\mu\in\mathscr{C}_K}\left(\sum_{s\in W}(\operatorname{sgn} s)p_n(s\mu-\lambda-\rho_n)\right)\right)\operatorname{ch}_K(\mu), \tag{3.6.3.16}$$

where

$$\operatorname{ch}_K(\lambda)=\sum_{s\in W}\frac{\operatorname{sgn}(s)\chi_s(\mu)}{D_K}$$

is the character of the representation of K, of highest weight $\mu-\rho_c$. Expression (3.6.3.16) leads us to suspect that the multiplicity of the representation of K with highest weight μ would occur in the discrete series representation π with multiplicity

$$\sum_{s\in W}(\operatorname{sgn} s)p_n(s(\mu+\rho_c)-\lambda-\rho_n). \tag{3.6.3.17}$$

This formula is indeed true **[HeSc1, Knap2, Wall2]**. It is known as Blattner's formula. A cohomological explanation and a generalization was found by Zuckerman in the context of his derived functor construction **[Knap1, Voga1, 2, Wall2]** (cf. §3.6.5).

(j) Study of the geometry of Weyl chambers shows that the norm (with respect to the Killing form) of $\lambda+\rho_n+\gamma$ is always greater than the norm of $\lambda+\rho_n$ if $\gamma\neq 0$. Hence the only term in the sum (3.6.3.15) which yields $\chi_{\lambda+\rho_n}$ is $\gamma=0$. It follows from Blattner's formula (3.6.3.17) that the representation of K with highest weight $\lambda+\rho_n-\rho_c=\lambda+\rho_{\mathscr{G}}-2\rho_c$ occurs in the discrete series π of formula (3.6.3.9) with multiplicity one. This K-representation is known as the *lowest* K*-type* of π. Vogan **[Voga2, 4]** has shown that every representation of G contains a K-type, which is lowest in a certain sense, with multiplicity one.

(k) Even more than in the Weyl character formula for compact groups, one sees the necessity for various "ρ shifts" to make parameters match up, particularly for the lowest K-type. To make sense of these, one should keep in mind that what is controlling everything is the infinitesimal character, which is computed by a ρ-shift coming from the Harish-Chandra homomorphism (cf. Theorem 3.5.5.23). In discussing lowest K-types, we must deal with the ρ-shifts for both K and G.

3.6.4. Classification. In the 1970s two classifications of the irreducible admissible representations of a semisimple Lie group were given, one by R. Langlands **[Lgld4]**, and another by D. Vogan **[Voga4]**. Somewhat earlier, a classification of the representations of complex groups had been given by Zhelobenko **[Zhel2]** and Parthasarathy-Rao-Varadarajan **[PaRV]**. Langlands' classification was similar in flavor. Although the classifications of Langlands and Vogan seem to be based on rather different principles—Langland's on the asymptotic behavior of matrix coefficients, and Vogan's on the behavior under restriction to the maximal compact subgroup K, in particular the existence of a "lowest K-type"—they were successfully combined by Vogan in [**Voga2**]. The result is a description of a standard realization of a given irreducible admissible representation, supplemented by information about its restriction to K.

In [**BeBe**], J. Bernstein and A. Beilinson announced a classification based on another circle of ideas. In particular it relies heavily on the theory of "D-modules"—modules for the sheaf of differential operators on a manifold. Despite its exotic origins, the Bernstein-Beilinson classification has a strong geometric flavor that gives it considerable appeal. The project of comparing and coordinating the features of the three classification schemes has been pursued in recent years by a group including H. Hecht, D. Milicic, W. Schmid, and J. Wolf [**HMSW**].

In this section we will give a brief description of the Langlands classification, which essentially describes representations in terms of standard embeddings in the principal series.

Let G be a semisimple Lie group, and let $P \subseteq G$ be a parabolic subgroup (cf. §A.2.4). Write $P = MAN$ as in formula (3.6.1.5). The abelian group A acts on the Lie algebra $\mathbf{n}$ of N by the conjugation action $\operatorname{Ad} A$. Under this action $\mathbf{n}$ decomposes into a direct sum of eigenspaces

$$\mathbf{n} = \sum_{\alpha} \mathbf{n}_{\alpha}, \qquad \alpha \in \Sigma^{+}(A, n), \tag{3.6.4.1}$$

where the α are the roots of A acting on $\mathbf{n}$:

$$\operatorname{Ad} a(n) = \alpha(a)n, \qquad a \in A,\ n \in \mathbf{n}_{\alpha}. \tag{3.6.4.2}$$

Evidently from their definition, the α are homomorphisms from A to $\mathbf{C}^{\times}$. In fact $\operatorname{Ad} A$ is a real-diagonalizable action, so that the α have images in $\mathbf{R}^{+\times}$. Write

$$A^{+} = \{a \in A : \alpha(a) \geq 1,\ \text{all } \alpha\}. \tag{3.6.4.3}$$

Note that A^{+} is a closed semigroup of A, with nonempty interior A^{+0}. Dually, write

$$(\widehat{A}^{\mathbf{C}})^{+} = \{\psi \in \widehat{A}^{\mathbf{C}} : |\psi(a)| > 1,\ \text{all } a \in A^{+0}\}. \tag{3.6.4.4}$$

Note that, in contrast to A^{+}, our definition makes $(\widehat{A}^{\mathbf{C}})^{+}$ an *open* semigroup in $\widehat{A}^{\mathbf{C}}$. This slight inconsistency will reduce by a little the total amount of notation that we need.

THEOREM 3.6.4.5 (Langlands classification). *Fix a minimal parabolic subgroup P_0 inside the connected semisimple Lie group G. Let P be a parabolic subgroup containing P_0, with Langlands decomposition $P = MAN$. Let σ be an irreducible tempered (in particular, unitary) representation of M, and $\psi \in (\widehat{A}^{\mathbf{C}})^{+}$. Then the principal series representation* P.S.(σ, ψ) *has a unique* *irreducible quotient* L.Q.(σ, ψ). *Each irreducible admissible representation of G is isomorphic to a unique* L.Q.(σ, ψ).

REMARKS. (a) For generic ψ, more specifically, for ψ not satisfying certain integrality conditions **[SpVo]**, the representations P.S.(σ, ψ) are irreducible, i.e., L.Q.$(\sigma, \psi) =$ P.S.(σ, ψ). However, the nature of L.Q.(σ, ψ) can change drastically as ψ varies. The difficulty of describing L.Q.(σ, ψ) more explicitly than is done in Theorem 3.6.4.5 is one reason why various outstanding problems, e.g., the classification of the unitary dual, remain unsolved. Thus while Theorem 3.6.4.5 gives us a place to put each representation, it does not provide us with a complete picture of the structure of the representations.

(b) A version of Theorem 3.6.4.5 first appeared in **[Lgld4]**. Its proof relied on results of an unpublished manuscript of Harish-Chandra on asymptotic expansions of matrix coefficients **[HaCh13]**. Harish-Chandra's results were refined and simplified by several authors; we refer in particular to **[CaMi]**. Relatively streamlined and complete accounts of these matters, as well as most of the rest of semisimple representation theory, are available in the texts **[Knap2]** and **[Wall2]**.

(c) In fact, the term "Langlands classification" for Theorem 3.6.4.5 is something of a misnomer. Langlands' goal in **[Lgld4]** was a different classification (see §4.2).

(d) In remark (i) following Theorem 3.6.3.7, we noted that if σ is a discrete series representation, then the restriction of σ to the maximal compact subgroup K contains a certain minimal K-type with multiplicity one. Vogan **[Voga4]** generalized the notion of minimal K-type to apply to any irreducible representation, and showed that his minimal K-type always occurs with multiplicity one. The K-module structure of the representations P.S.(σ, ψ) is independent of ψ, so they all have the same minimal K-type μ_σ. Vogan shows **[Voga2]** that μ_σ survives in L.Q.(σ, ψ). This additional information about L.Q.(σ, ψ) allows one to characterize it as a subquotient of a broader class of principal series representations.

EXAMPLE. Parametrize the spherical principal series of $\mathrm{SL}_2(\mathbf{R})$ by $\lambda \in \mathbf{C}$, as in formula (3.6.1.8). Then the unitary principal series, which are all tempered, lie on the imaginary axis. They are all irreducible, and the representations labeled by λ and by $-\lambda$ are mutually equivalent. The representations associated to P_0 in Theorem 3.6.4.5 are those in the right half plane. All are irreducible, except for $\lambda = 1, 3, 5, 7, \ldots$. For these, the Langlands quotient is the finite-dimensional representation of dimension λ (cf. §3.1). The remaining constituents (there are two) of P.S.$(1, 2m+1)$ are discrete series, hence get counted among the tempered representations. The story is similar for the nonunitary principal series P.S.$(\tilde{\varepsilon}, \tilde{a}^\lambda)$ ($\tilde{\varepsilon}$ here being the nontrivial character of M_0; cf. formula (3.6.1.8)), except that P.S.$(\tilde{\varepsilon}, 1)$, the point of symmetry of the unitary principal series, reduces into two pieces, and the points of reducibility in the right half plane are at the even integers $\lambda = 2, 4, 6, 8, \ldots$, the Langlands quotient again being the finite-dimensional representation of dimension λ.

The representations $\mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{\lambda})$ for $\mathrm{Re}\,\lambda < 0$ are contragredient to $\mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{-\lambda})$ (cf. §A.1.10). If $\mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{\lambda})$ is irreducible, it is equivalent to $\mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{-\lambda})$. If it is reducible, it has the same constituents as $\mathrm{P.S.}(\tilde{\varepsilon}, \tilde{a}^{-\lambda})$, but the finite-dimensional representation is now a subrepresentation, and the two discrete series are quotients.

3.6.5. Derived functor modules. Part of the fascination of semisimple harmonic analysis is the strong interaction of algebra and analysis that it affords. After his initial, heavily algebraic, papers on foundational issues in representation theory, Harish-Chandra's methods became more and more analytic, culminating in the construction of the discrete series via a deep study of character theory [**HaCh13-20**]. Then, beginning in the late 1960s, the work of Dixmier [**Dixm1**], Kostant [**Kost8**], and the Gelfand school (cf. [**BGG1–3**]) reemphasized the algebraic aspects of the theory.

In the late 1960s, there had been efforts, by W. Schmid [**Schm1–3**], and others [**OkOz, Hott, Part1**], to "realize the discrete series," i.e., construct specific vector spaces on which a semisimple group could act in a natural way, and such that the resulting representation of G was a given discrete series representation (cf. §3.6.3). G. Zuckerman, considering Schmid's work from the more algebraic point of view, invented a flexible and fruitful purely algebraic method for constructing representations. Parthasarathy [**Part2**] had an idea for a similar construction, and Enright-Varadarajan [**EnVa**] also proposed an interesting, though conceptually less transparent, algebraic method for constructing representations. After several simplifications and refinements [**Voga1, EnWa, DuVe2, KnVo, Wign2, Wall2**], Zuckerman's method, which has become known as the "derived functor construction," has become a standard tool for constructing representations. In particular, it can be used to give an a priori construction of the discrete series [**Wall2**]. (Proof of "exhaustion"—that all discrete series are so realized—still, however, requires analysis, especially character theory.) It also yields other interesting classes of nontempered unitary representations, e.g., the representations with nontrivial $(\mathbf{g}, \mathbf{k})$-cohomology [**VoZu**]. We will give a brief description of Zuckerman's idea. A nice overview is given in [**Voga1**], a leisurely introductory treatment is in [**Knap2**], and detailed expositions are given in [**Voga2**] and [**Wall2**].

The inspiration behind the derived functor construction is epistemologically interesting: it consists in taking seriously what might have seemed merely a technical convenience—the notion of $(\mathbf{g}, K)$-module. The general goal of the algebraic approach is to replace a G-module by a suitable $\mathscr{U}(\mathbf{g})$-module which can serve as its proxy, i.e., that will mirror its essential features. A first guess might be the subspace of smooth vectors (cf. §A.1.13) but this is of uncountable dimension, hence too large to be studied algebraically. At the outset, however, Harish-Chandra had shown **[HaCh2]** that if ρ is a t.c.i. (§A.1.7) (or quasisimple) irreducible representation of a semisimple group G with maximal compact subgroup K on a Banach space V, then the subspace V_K of K-finite vectors (vectors contained in finite-dimensional, K-invariant subspaces) is invariant under the action of the Lie algebra $\mathbf{g}$. Hence V_K is a module for $\mathbf{g}$. It is also obviously a module for K, and the two module structures are compatible in some obvious ways:

(3.6.5.1) (i) The differential of the action of K is obtained by restriction of the $\mathbf{g}$-action to the Lie subalgebra $\mathbf{k} \subseteq \mathbf{g}$ corresponding to K.
(ii) K normalizes $\mathbf{g}$ inside End V_K, and conjugation by K in End V_K yields the usual adjoint action Ad K on $\mathbf{g}$.

These properties were enshrined by Gelfand **[Gelf]** in a formal definition of what is now usually called a $(\mathbf{g}, K)$*-module.* Our space V_K also has the property that each K-isotypic subspace V_μ, $\mu \in \widehat{K}$, is finite dimensional. A $(\mathbf{g}, K)$-module with this property is called an *admissible* $(\mathbf{g}, K)$-module or *Harish-Chandra* module.

For a given irreducible representation of G on V, the associated Harish-Chandra module V_K is convenient to work with. For example, for $\mathrm{SL}_2(\mathbf{R})$ it is an easy, pleasant exercise to determine all possible irreducible $(\mathbf{sl}_2(\mathbf{R}), \mathrm{SO}_2)$ modules, and to check which ones could carry an invariant inner product. This was the method of Bargmann **[Barg1]** at the very beginning of semisimple representation theory. Further, V_K captures much of what we want to know about V: for example, Harish-Chandra showed **[HaCh3]** that V could be unitary if and only if V_K carries an invariant positive-definite Hermitian form, and, in that case, V_K determines V up to unitary equivalence.

Thus, one might propose Harish-Chandra modules as a technically convenient class of $\mathbf{g}$-modules with which to work in order to study algebraic aspects of representation theory. However, $(\mathbf{g}, K)$-modules also appear to be rather awkward from certain points of view. Some people might be put off by the loss of symmetry entailed by the choice of a particular maximal compact group K. A serious technical problem is presented by induction. Induction is a basic method for constructing representations, so we would like to have an algebraic version of it. There is in fact a standard notion of induction in the category of associative algebras, defined in terms of tensor products. Thus suppose $\mathbf{g}$ is a Lie algebra, $\mathbf{h} \subseteq \mathbf{g}$ is a subalgebra, and V is an $\mathbf{h}$-module. Define

$$\mathrm{ind}_{\mathbf{h}}^{\mathbf{g}} V = \mathscr{U}(\mathbf{g}) \otimes_{\mathscr{U}(\mathbf{h})} V. \tag{3.6.5.2}$$

We recall **[Jaco2]** that $\mathscr{U}(\mathbf{g}) \otimes_{\mathscr{U}(\mathbf{h})} V$ is the quotient of the usual tensor product $\mathscr{U}(\mathbf{g}) \otimes V$ by the subspace spanned by tensors of the form $xy \otimes v - x \otimes y(v)$, $x \in \mathscr{U}(\mathbf{g})$, $y \in \mathscr{U}(\mathbf{h})$, $v \in V$. (In fact, it suffices to take $y \in \mathbf{h}$.) The action of $\mathscr{U}(\mathbf{g})$ on $\mathrm{ind}_{\mathbf{h}}^{\mathbf{g}} V$ is the push-down of left multiplication on $\mathscr{U}(\mathbf{g})$.

A variant notion, more or less dual to (3.6.5.2), also exists. It is sometimes called *production*, (though some feel the terminology for induction and production should be reversed) and is defined by

$$\text{(3.6.5.3)} \qquad \mathrm{pro}_{\mathbf{h}}^{\mathbf{g}} V = \mathrm{Hom}_{\mathbf{h}}(\mathscr{U}(\mathbf{g}), V).$$

Here the action of $\mathbf{h}$ on $\mathscr{U}(\mathbf{g})$ is via multiplication by $\mathbf{h}$ on the left. The action of $\mathbf{g}$ on $\mathrm{pro}_{\mathbf{h}}^{\mathbf{g}}(V)$ is via multiplication on the right in $\mathscr{U}(\mathbf{g})$. If M is a subgroup of G, whose Lie algebra $\mathbf{m}$ is contained in $\mathbf{h}$, and V is an $(\mathbf{h}, M)$-module, we let M act on $\mathrm{pro}_{\mathbf{h}}^{\mathbf{g}} V$ by the recipe

$$m(f)(u) = m(f(\mathrm{Ad}\, m(u))), \qquad m \in M, f \in \mathrm{pro}_{\mathbf{h}}^{\mathbf{g}} V, u \in \mathscr{U}(\mathbf{g}).$$

We then replace $\mathrm{pro}_{\mathbf{h}}^{\mathbf{g}} V$ by the subspace of its M-finite vectors and so obtain a $(\mathbf{g}, M)$-module.

Both induction **[Wall2, Knap1]** and production **[Voga1, 2, Knap1]**, are used in accounts of derived functor modules. Other conventions concerning the derived functor construction also vary from author to author, necessitating an annoying, if in principle straightforward, translation process to compare results. We follow the conventions of **[Voga1]**; in particular, we use production.

The production process (3.6.5.3) converts $\mathbf{h}$-modules to $\mathbf{g}$-modules, but it is unlikely that it will yield modules which are spanned by K-finite (or, what is more reasonable to discuss at this stage, $\mathbf{k}$-finite) vectors, or which even contain any $\mathbf{k}$-finite vectors at all. This would seem to be a serious drawback of the $(\mathbf{g}, K)$-module formalism. However, Zuckerman saw how to make a virtue of necessity, and converted this seeming liability into a construction method that is subtler than production, but still is fairly manageable.

The most obvious way to associate a $(\mathbf{g}, K)$-module to a $\mathbf{g}$-module is to look at the submodule of $\mathscr{U}(\mathbf{k})$-finite vectors, then exponentiate to get an action of K on this submodule. (There is further fussing necessary if K is not connected and simply connected **[Voga1, 2, Knap1, Wall2]**. We will ignore this fussing. Thus we will actually be discussing $(\mathbf{g}, \mathbf{k})$-modules (the definition of which is hoped to be obvious) rather than $(\mathbf{g}, K)$-modules.) This procedure, however, may well result in the trivial module. Zuckerman observed, however, that the process of passing from a $\mathbf{g}$-module to the $(\mathbf{g}, \mathbf{k})$-module of $\mathbf{k}$-finite vectors is a functor, and moreover, it is a left-exact functor. Thus we have the possibility of taking its (right) derived functors; and even if the module we start with has no K-finite vectors, one of the higher derived functors may be nontrivial. This does in fact happen in interesting cases.

Thus if V is a $\mathbf{g}$-module, we can define

$$\text{(3.6.5.4)} \qquad \Gamma(V) = \{v \in V : \dim \mathscr{U}(\mathbf{k})(v) < \infty\}.$$

(As noted above, one must give a slightly more complicated definition of Γ if K is not connected or not simply connected.) The derived functors of Γ

will be denoted Γ^i. A curious point about Γ is that, while Γ is a functor on $\mathbf{g}$-modules, it depends only on the $\mathbf{k}$-module structure of these modules. Further, we can express Γ in terms of $\mathbf{k}$-fixed vectors in certain auxiliary modules, viz.

$$\Gamma(V) = \sum_{\sigma\in\hat{\mathbf{k}}} (V \otimes \sigma^*)^{\mathbf{k}} \otimes \sigma. \tag{3.6.5.5}$$

Here $\hat{\mathbf{k}}$ is the collection of irreducible finite-dimensional representations of $\mathbf{k}$—the same as $\widehat{K}$ if K is connected and simply connected—and $(V\otimes\sigma^*)^{\mathbf{k}}$ indicates the $\mathbf{k}$-invariant vectors in $V\otimes\sigma^*$.

Since formula (3.6.5.5) expresses $\Gamma(V)$ in terms of the functor of K-fixed vectors, it suggests the Γ^i should be expressible in terms of the derived functors of the $\mathbf{k}$-fixed vector functor, which is Lie algebra cohomology [**BoWa**, **Jaco1**, **Knap1**]. This is not immediate, since the construction of derived functors depends on injective resolutions, and the notion of injectiveness depends on the category in which one is working. However, we can construct injective (or projective) resolutions of $\mathbf{g}$-modules with injective (or projective) modules of the form $\mathrm{Hom}(\mathscr{U}(\mathbf{g}), Y)$ (or $\mathscr{U}(\mathbf{g})\otimes Y$), where Y is a $\mathbf{g}$-module. Since $\mathscr{U}(\mathbf{g})$ is free as a $\mathscr{U}(\mathbf{k})$-module, by the Poincaré-Birkhoff-Witt Theorem [**Jaco1**, **Serr2**, **Hump**], one can check that these are injective (or projective) as $\mathbf{k}$-modules also. It follows that

$$\Gamma^i(V) \simeq \sum_{\sigma\in\hat{\mathbf{k}}} H^i(\mathbf{k}, V\otimes\sigma^*)\otimes\sigma, \tag{3.6.5.6}$$

where $H^i(\mathbf{k}, X)$ is the ith Lie algebra cohomology of the $\mathbf{k}$-module X.

There is also a relative version of (3.6.5.6). Suppose we are given a subalgebra $\mathbf{m}\subseteq\mathbf{k}$, and we start with a $\mathbf{g}$-module V which is already $\mathbf{m}$-finite (a $(\mathbf{g}, \mathbf{m})$-module). Then it is appropriate to work inside the category of $\mathbf{m}$-finite $\mathbf{g}$-modules, and the relevant derived functors of the $\mathbf{k}$-fixed vector functor are the relative Lie algebra cohomology groups [**BoWa**, **Knap1**]. Thus, if V is an $\mathbf{m}$-finite $\mathbf{g}$-module, or a $(\mathbf{g}, M)$-module for some $\mathbf{m}\subseteq\mathbf{k}$, then

$$\Gamma^i(V) \simeq \sum_{\sigma\in\hat{\mathbf{k}}} H^i(\mathbf{k}, \mathbf{m}, V\otimes\sigma^*)\otimes\sigma, \tag{3.6.5.7}$$

where $H^i(\mathbf{k}, \mathbf{m}, X)$ is the ith $(\mathbf{k}, \mathbf{m})$-relative cohomology group of the $\mathbf{m}$-finite $\mathbf{k}$-module X.

This describes the structure of $\Gamma^i(V)$ as a $\mathbf{k}$-module, but of course our interest in it is as a $\mathbf{g}$-module. A simple observation [**EnWa**, **DuVe**, **KnVo**, **Wign2**] allows this to be done simply. Note that, if V is a $\mathbf{k}$-module, and X is a finite-dimensional $\mathbf{k}$-module, then

$$\Gamma(X\otimes V) \simeq X\otimes\Gamma(V). \tag{3.6.5.8}$$

In fact, this isomorphism is a natural equivalence of functors. Further, tensoring with X is exact. It follows easily that

$$\Gamma^i(X\otimes V) \simeq X\otimes\Gamma^i(V), \tag{3.6.5.9}$$

and that this also is an equivalence of functors, i.e., this isomorphism is natural.

Now suppose V is a $\mathbf{g}$-module. This means we have a mapping

$$\begin{aligned}(3.6.5.10)\qquad &\mathbf{g}\otimes V \xrightarrow{\mu} V,\\ &\mu(x\otimes v) = x(v), \qquad x\in\mathbf{g},\, v\in V.\end{aligned}$$

Clearly the mapping μ of (3.6.5.10) determines the $\mathbf{g}$-module structure of V. Furthermore, the fact that μ defines a $\mathbf{g}$-module structure on V can be expressed solely in terms of μ: we should have the identity

$$(3.6.5.11)\qquad \mu(x_1\otimes\mu(x_2\otimes v)) - \mu(x_2\otimes\mu(x_1\otimes v)) = \mu([x_1, x_2]\otimes v)$$

as mappings from $\mathbf{g}\otimes\mathbf{g}\otimes V$ to V. Combining these remarks with the previous paragraph, we see that the $\mathbf{g}$-module structure (3.6.5.10) on V induces a $\mathbf{g}$-module structure on $\Gamma^i(V)$ for all i. It is very plausible that this $\mathbf{g}$-module structure is the one that should be carried by $\Gamma^i(V)$, and it is indeed so (cf. references preceding (3.6.5.8)).

Now let us introduce the modules to which we wish to apply the Γ^i. Consider $x\in\mathbf{k}$. Let $\mathbf{l}$ be the centralizer of x in $\mathbf{g}$. The complexification $\mathbf{l}_{\mathbf{C}}$ is a Levi component of a parabolic subalgebra of $\mathbf{g}_{\mathbf{C}}$. More precisely, since x is in $\mathbf{k}$, $\operatorname{ad} x$ will act on $\mathbf{g}_{\mathbf{C}}$ with purely imaginary eigenvalues. Let $\mathbf{n}^+\subseteq\mathbf{g}_{\mathbf{C}}$ be the sum of the $\operatorname{ad} x$ eigenspaces with eigenvalues with positive imaginary part, and let $\mathbf{n}^-$ be the complex conjugate of $\mathbf{n}^+$, the sum of $\operatorname{ad} x$-eigenspaces whose eigenvalues have negative imaginary part. Then

$$(3.6.5.12)\qquad \mathbf{g}_{\mathbf{C}} = \mathbf{n}^-\oplus\mathbf{l}_{\mathbf{C}}\oplus\mathbf{n}^+$$

and $\mathbf{q} = \mathbf{l}_{\mathbf{C}}\oplus\mathbf{n}^+$ is a parabolic subgroup of $\mathbf{g}_{\mathbf{C}}$. Let θ be the Cartan involution on $\mathbf{g}$: the automorphism of order 2 whose fixed point set is $\mathbf{k}$ (cf. §A.2). Since $x\in\mathbf{k}$, it is clear that $\mathbf{l}$ and $\mathbf{n}^\pm$ are invariant under θ. The parabolic subalgebra $\mathbf{q}$ is therefore called a *θ-stable parabolic* subalgebra.

Let $\mathbf{g} = \mathbf{l}_{\mathbf{C}}\oplus\mathbf{n}^+$ be the θ-stable parabolic subalgebra defined by $x\in\mathbf{k}$, as above. Let $m = \dim\mathbf{n}^+$. The adjoint action of $\mathbf{l}_{\mathbf{C}}$ on $\mathbf{n}^+$ gives rise to a one-dimensional action of $\mathbf{l}_{\mathbf{C}}$ on $\Lambda^m(\mathbf{n}^+)$, the top exterior power of $\mathbf{n}^+$. Denote by $2\rho_{\mathbf{q}}$ the weight in $\mathbf{l}_{\mathbf{C}}^*$ defined by $\Lambda^m(\mathbf{n}^+)$. Note that $2\rho_{\mathbf{q}}$ is a formal analog of the modular function of a real parabolic subgroup (cf. §A.1.15 and formula (3.5.5.16)) and we use it in the same way: to normalize induction by twisting with ρ before induction (cf. equation (3.6.1.3)). Let $\mathbf{C}_{\rho_{\mathbf{q}}}$ be a one-dimensional module on which $\mathbf{l}_{\mathbf{C}}$ acts by the weight ρ. (In some sense, $\mathbf{C}_{\rho_{\mathbf{q}}} =$ "$(\Lambda^m(\mathbf{n}^+))^{1/2}$.")

Let Z be an $(\mathbf{l}, \mathbf{l}\cap\mathbf{k})$-module. Extend by complex linearity the action of $\mathbf{l}$ to $\mathbf{l}_{\mathbf{C}}$. Extend the action of $\mathbf{l}_{\mathbf{C}}$ to an action of $\mathbf{q}$ by letting $\mathbf{n}^+$ act trivially. Define

$$(3.6.5.13)\qquad \mathscr{R}^j(Z) = \Gamma^j(\operatorname{pro}_{\mathbf{q}}^{\mathbf{g}_{\mathbf{C}}}(Z\otimes\mathbf{C}_{\rho_{\mathbf{q}}})_{(\mathbf{k}\cap l)}), \qquad j\geq 0.$$

Here the sub-$(\mathbf{k}\cap l)$ means the submodule of $(\mathbf{k}\cap l)$-finite vectors. The $\mathscr{R}^j$ are functors which transform $(\mathbf{l}, \mathbf{l}\cap\mathbf{k})$-modules into $(\mathbf{g}, \mathbf{k})$-modules (or, more carefully done, $(\mathbf{l}, L\cap K)$-modules into $(\mathbf{g}, K)$-modules). The $\mathscr{R}^j(Z)$ are the *derived functor modules*, and the functors $\mathscr{R}^j$ are often referred to as *cohomological induction*. The usual yoga about derived functors **[Jaco2, Lang3]** shows that if

(3.6.5.14a) $$0 \to Z' \to Z \to Z'' \to 0$$

is a short exact sequence of $(\mathbf{l}, \mathbf{l}\cap\mathbf{k})$-modules, then the $\mathscr{R}^j(Z)$ can be organized into a long exact sequence

(3.6.5.14b) $$\to \mathscr{R}^j(Z) \to \mathscr{R}^j(Z'') \to \mathscr{R}^{j+1}(Z') \to \mathscr{R}^{j+1}(Z) \to .$$

Further, the $\mathscr{R}^j$ are compatible with the Harish-Chandra homomorphism (cf. §3.5.5). Suppose Z has an infinitesimal character; that is, suppose there is a homomorphism

$$\lambda : \mathscr{Z}\mathscr{U}(\mathbf{l}) \to \mathbf{C}$$

of the center of the enveloping algebras of $\mathbf{l}$, such that $u(z) = \lambda(u)z$ for $u \in \mathscr{Z}\mathscr{U}(\mathbf{l})$ and $z \in Z$; or, in other words, $\mathscr{Z}\mathscr{U}(\mathbf{l})$ acts on Z by scalars. Then the $\mathscr{R}^j(Z)$ will also have an infinitesimal character (independent of j), and this character is determined by the Harish-Chandra homomorphism (cf. Theorem 3.5.5.23). Precisely, if $\tilde{p} : \mathscr{Z}\mathscr{U}(\mathbf{g}) \to \mathscr{Z}\mathscr{U}(\mathbf{l})$ is the Harish-Chandra homomorphism, then the infinitesimal character of the $\mathscr{R}^j(Z)$ is $\lambda \circ \tilde{p}$. This is obvious for $\mathscr{R}^0(Z)$, and follows for $j > 0$ by an argument similar to that used in the current proofs of Theorem 3.5.5.20.

One can also show that $\mathscr{R}^j$ takes finite-length modules to finite-length modules, and Harish-Chandra (i.e., admissible) modules to Harish-Chandra modules.

The most interesting question, of course, is when are the $\mathscr{R}^j(Z)$ nonzero. Because the $\mathscr{R}^j$ are computed, at least as $\mathbf{k}$-modules, in terms of relative $(\mathbf{k}, \mathbf{l}\cap\mathbf{k})$-cohomology as per equation (3.6.5.7), and since the standard complex **[BoWa, Knap1]** for computing relative cohomology has length $\dim(\mathbf{k}/\mathbf{k}\cap\mathbf{l})$, a first conclusion is that $\mathscr{R}^j(Z) = 0$ if $j > \dim(\mathbf{k}/(\mathbf{k}\cap\mathbf{l}))$. However, a stronger result holds. Observe that

$$(\mathbf{k}/(\mathbf{k}\cap\mathbf{l}))_{\mathbf{C}} \simeq (\mathbf{k}_{\mathbf{C}}\cap\mathbf{n}^+) \oplus (\mathbf{k}_{\mathbf{C}}\cap\mathbf{n}^-).$$

Thus

$$\dim(\mathbf{k}/(\mathbf{k}\cap\mathbf{l})) = 2\dim(\mathbf{k}_{\mathbf{C}}\cap\mathbf{n}^-) = 2\dim(\mathbf{k}_{\mathbf{C}}/(\mathbf{k}_{\mathbf{C}}\cap\mathbf{q})).$$

The modules we are dealing with are defined in terms of the production constructions as described in equation (3.6.5.3). For produced modules, one can construct a special resolution (see **[Wall2, §6.A.14]** for the analog for induction) which implies they have relative cohomology only in a restricted range. Precisely, the relative $(\mathbf{k}_{\mathbf{C}}, (\mathbf{k}\cap\mathbf{l})_{\mathbf{C}})$-cohomology of a $(\mathbf{k}_{\mathbf{C}}, (\mathbf{k}\cap\mathbf{l})_{\mathbf{C}})$-module produced from $\mathbf{k}_{\mathbf{C}}\cap\mathbf{q}$ will vanish in degrees above

$$\dim \mathbf{k}_{\mathbf{C}}/(\mathbf{k}_{\mathbf{C}}\cap\mathbf{q}) = \tfrac{1}{2}\dim(\mathbf{k}/\mathbf{k}\cap\mathbf{l}).$$

The modules with which we are working have the form $\mathrm{pro}_{\mathbf{q}}^{\mathbf{g}_{\mathbf{C}}}(Z \otimes \mathbf{C}_{\rho_{\mathbf{q}}})$. Their restrictions to $\mathbf{k}$ are not precisely of the form $\mathrm{pro}_{\mathbf{k}_{\mathbf{C}} \cap \mathbf{q}}^{\mathbf{k}} Y$, but they have a filtration by submodules whose quotients are of this form, and this is sufficient to establish the vanishing.

THEOREM 3.6.5.15. *$\mathscr{R}^j(Z) = 0$ for $j \geq \dim(\mathbf{k}_{\mathbf{C}}/(\mathbf{k}_{\mathbf{C}} \cap \mathbf{q}))$.*

On the other hand, relative Lie algebra cohomology features a version of Poincaré duality [**BoWa**, **KnVo**, **Wall2**]. This allows one to show, under a certain positivity condition [**Knap1**, **Voga1**, **Wall2**] on the infinitesimal character of Z, that $\mathscr{R}^j(Z) = 0$ also for $j < \dim(\mathbf{k}_{\mathbf{C}}/(\mathbf{k}_{\mathbf{C}} \cap \mathbf{q}))$. Thus, under these hypotheses, we have $\mathscr{R}^j(Z) = 0$ except for $j = S = \frac{1}{2}\dim(\mathbf{k}/(\mathbf{k} \cap \mathbf{l}))$. The function, roughly, of the positivity condition is to make the ($(\mathbf{k} \cap \mathbf{l})$-finite vectors in the) produced modules $\mathrm{pro}_{\mathbf{q}}^{\mathbf{g}_{\mathbf{C}}}(Z \otimes \mathbf{C}_{\rho_{\mathbf{q}}})$ irreducible, or at least to guarantee that these modules carry nondegenerate Hermitian forms (the Shapovalov form [**Voga1, 3**]); this then guarantees that their Hermitian dual modules are again produced modules of the same sort, so that Theorem 3.6.5.15 applies to them too. Then Poincaré duality, which relates the Hermitian dual of $\mathscr{R}^j(Z)$ to Γ^{2s-j} of a module constructed from the Hermitian dual of Z, guarantees vanishing of $\mathscr{R}^j$ except for $j = S$.

On the other hand, it is clear that some kind of condition on Z is needed to guarantee vanishing of $\mathscr{R}^j(Z)$ for $j < S$. For example, in the case when $\mathbf{l} = \mathbf{k} \cap \mathbf{l} = \mathbf{t}$ is a Cartan subalgebra, and $Z = \mathbf{C}_\mu$ is one dimensional we see that $\mathrm{pro}(Z)$ is the dual of a Verma module (cf. equation (3.6.5.21) below), which will contain a finite-dimensional representation if μ is negative.

REMARK. It may be instructive to make a few elementary observations about the structure of $\mathrm{pro}_{\mathbf{q}}^{\mathbf{g}_{\mathbf{C}}}(Z)_{(\mathbf{k}\cap l)}$. For this digression, we will abbreviate

$$\mathrm{pro}_{\mathbf{q}}^{\mathbf{g}_{\mathbf{C}}}(Z)_{(\mathbf{k}\cap l)} = \mathrm{pro}(Z).$$

By the Poincare-Birkhoff-Witt Theorem [**Jaco1**, **Serr2**, **Hump**], multiplication on $\mathscr{U}(\mathbf{g})$ induces a linear isomorphism

$$\mathscr{U}(\mathbf{g}) \simeq \mathscr{U}(\mathbf{n}^-) \otimes \mathscr{U}(\mathbf{q}). \tag{3.6.5.16}$$

It follows that a mapping $T \in \mathrm{Hom}_{\mathbf{q}}(\mathscr{U}(\mathbf{g}), Z)$ is determined by its values on $\mathscr{U}(\mathbf{n}^-)$; and conversely, any mapping from $\mathscr{U}(\mathbf{n}^-)$ to Z extends to a $\mathbf{q}$-module map from $\mathscr{U}(\mathbf{g})$ to Z. Thus, as linear spaces, we have the isomorphism

$$\mathrm{pro}(Z) \simeq \mathrm{Hom}(\mathscr{U}(\mathbf{n}^-), Z)_{(\mathbf{k}\cap l)}. \tag{3.6.5.17}$$

Consider the action of $\mathbf{l}$ on the module. The definition of the $\mathbf{g}$-action is by multiplication on the right:

$$l(T)(u) = T(ul), \qquad l \in \mathbf{l},\ T \in \mathrm{Hom}_{\mathbf{q}}(\mathscr{U}(\mathbf{g}), Z).$$

On the other hand, the condition that T be a $\mathbf{q}$-module map is that

$$T(lu) = l(T(u)).$$

Thus we may write

$$\begin{aligned} l(T)(u) &= T(ul) - T(lu) + l(T(u)) \\ &= -T([l, u]) + l(T(u)). \end{aligned}$$

Since $\mathbf{u}^-$ is stable under ad $\mathbf{l}$, this formula allows us to see the linear isomorphism (3.6.5.17) as an isomorphism of $\mathbf{l}$-modules, if we define

(3.6.5.18)

$$l(T_1)(u) = -T_1(\text{ad}\, l(u)) + l(T_1(u)),$$
$$l \in \mathbf{l},\ u \in \mathscr{U}(\mathbf{n}^-),\quad T_1 \in \text{Hom}(\mathscr{U}(\mathbf{n}^-), Z).$$

Note that this is the standard action of $\mathbf{l}$ on $\text{Hom}(\mathscr{U}(\mathbf{n}^-), Z)$, constructed from the actions ad on $\mathscr{U}(\mathbf{n}^-)$, and the given action on Z [**Serr2**, **Jaco1**, **Hump**].

Denote by $\mathbf{c}$ the center of $\mathbf{l}$; thus

(3.6.5.19) $$\mathbf{l} = \mathbf{c} \oplus [\mathbf{l}, \mathbf{l}],$$

where $[\mathbf{l}, \mathbf{l}]$ indicates the commutator ideal in $\mathbf{l}$. Note that $\mathbf{c} \subseteq \mathbf{k}$. Let $\Sigma_{\mathbf{n}^-} \subseteq (\mathbf{c}_{\mathbf{C}})^*$ be the set of weights for the adjoint action of $\mathbf{c}$ on $\mathbf{n}^-$. Then the weights of $\mathbf{c}$ acting on $\mathscr{U}(\mathbf{n}^-)$ have the form $\sum n_\beta \beta$, with β running through $\Sigma_{\mathbf{n}^-}$ and $n_\beta \in \mathbf{Z}^+$, the nonnegative integers. These all lie inside some proper cone in $i\mathbf{c}^*$. In particular, the multiplicity of any given weight for $\mathbf{c}$ in $\mathscr{U}(\mathbf{n}^-)$ is finite. This multiplicity is known as the *partition function.* We denote it by $p_{\mathbf{n}^-}$. Also, denote by $C_{\mathbf{n}^-}$ the set of all weights of $\mathbf{c}$ on $\mathscr{U}(\mathbf{n}^-)$. Write

(3.6.5.20) $$\mathscr{U}(\mathbf{n}^-) = \sum_{\gamma \in C_{\mathbf{n}^-}} \mathscr{U}(\mathbf{n}^-)_\gamma,$$

where $\mathscr{U}(\mathbf{n}^-)_\gamma$ is the $\mathbf{c}$-eigenspace for the weight γ. Note that $\dim \mathscr{U}(\mathbf{n}^-)_\gamma = p_{\mathbf{n}^-}(\gamma)$.

Suppose that $\mathbf{c}$ acts on the $(\mathbf{l}, \mathbf{k} \cap \mathbf{l})$-module Z by scalars; that is, Z consists of a single weight space for $\mathbf{c}$. Since this is automatically true if Z has an infinitesimal character, and since any $(\mathbf{l}, \mathbf{k} \cap \mathbf{l})$-module is a direct sum of its $\mathbf{c}$-weight spaces, the assumption that Z is a single weight space is only a mild restriction on Z. Note also that if Z is a weight space for $\mathbf{c}$, so is $Z \otimes \mathbf{C}_{\rho_{\mathbf{q}}}$. If the weight defined by Z is $\mu_Z = \mu \in (\mathbf{c}_{\mathbf{C}})^*$, then the weight defined by $Z \otimes C_{\rho_{\mathbf{q}}}$ is $\mu + \rho_{\mathbf{q}}$. A very important special case is when $Z = \mathbf{C}_\mu$, the one-dimensional module with weight μ.

By means of the decomposition (3.6.5.20) above, we may regard $\text{Hom}(\mathscr{U}(\mathbf{n}^-)_\gamma, Z)$ as a subspace of $\text{Hom}(\mathscr{U}(\mathbf{n}^-), Z)$. Under the hypotheses of the previous paragraph, it is a weight space for $\mathbf{c}$, with weight $\mu - \gamma$. It follows that $\text{pro}(Z)$ consists simply of the direct sum

(3.6.5.21) $$\text{pro}(Z) \simeq \sum_{\gamma \in C_{\mathbf{n}^-}} \text{Hom}(\mathscr{U}(\mathbf{n}^-)_\gamma, Z),$$

and the c-weights of this module are the set $\mu - C_{\mathbf{n}^-}$. In the case when $\mathbf{l} = \mathbf{k} \cap \mathbf{l} = \mathbf{t}$ is a Cartan subalgebra, we see that, as a $\mathbf{t}$-module, $\operatorname{pro}(Z)$ looks like a Verma module (cf. §3.5.3). However, it is not; the induced module (3.6.5.2) is a Verma module, and our $\operatorname{pro}(Z)$ is dual to a Verma module. But if the positivity condition mentioned above is satisfied, then the Verma module dual to $\operatorname{pro}(Z)$ is irreducible; hence, V itself is an irreducible Verma module, and one is in a position to use the vanishing Theorem 3.6.5.15.

In light of the discussion above, the question to focus on is, what is the structure of $\mathscr{R}^S(Z)$? When all $\mathscr{R}^j(Z)$ except $\mathscr{R}^S(Z)$ vanish, the Euler-Poincaré principle of cohomology [**Lang3**, p. 124] allows one to calculate the K-structure of $\mathscr{R}^S(Z)$. The result is a formula for the multiplicities of K-types in $\mathscr{R}^S(Z)$ in terms of an alternating sum over the Weyl group of the values of a partition function, not the partition function of all of $\mathbf{n}^-$, but of the noncompact part of $\mathbf{n}^-$. This is a generalization of the Blattner formula (cf. Remark (i) following Theorem 3.6.3.7). Despite its elegance, it is difficult to get specific information about general K-types from it. However, if Z is a character, then under the same positivity conditions that imply vanishing of $\mathscr{R}^j(Z)$ except for $j = S$, one can find one particular K-type, the analog of the lowest K-type for the discrete series (cf. Remark (i) again); in fact it is the lowest K-type in the sense of Vogan [**Voga2**, **4**], which occurs with multiplicity one. This implies in particular that $\mathscr{R}^S(Z) \neq 0$.

Thus for certain $(\mathbf{l}, \mathbf{l} \cap \mathbf{k})$-modules Z, we obtain a nontrivial $(\mathbf{g}, \mathbf{k})$-module $\mathscr{R}^S(Z)$, with (in principle) known K-structure. Furthermore, again under the same positivity hypotheses, Poincaré duality considerations enable one to show that if Z has an invariant Hermitian inner product, then so does $\mathscr{R}^S(Z)$, and in fact [**Voga3**, **Wall5**], if Z is unitary (i.e., the invariant Hermitian structure is positive definite), then so is $\mathscr{R}^S(Z)$. Hence one has a construction of unitary representations.

Among these representations are the discrete series. Again for simplicity, we assume that G, hence K and T, is connected, and we use the notation of §3.6.3.

THEOREM 3.6.5.22. *Let $\mathbf{l} = \mathbf{k} \cap \mathbf{l} = \mathbf{t}$ be a Cartan subalgebra of $\mathbf{g}$, and let $\mathbf{b} \supseteq \mathbf{t}$ be a (θ-stable) Borel subalgebra containing $\mathbf{t}$. Let T be the torus associated to $\mathbf{t}$. Choose $\mu \in \widehat{T} + \rho$, μ dominant for $\mathbf{b}$. Then $\mathscr{R}^S(\mathbf{C}_\mu)$ is the discrete series representation attached to μ.*

In Zuckerman's original formulation, this result was simply a recognition theorem, based on a characterization by Schmid [**Schm2**] of discrete series in terms of their K-spectrum and Casimir eigenvalues. However, subsequent work has enabled irreducibility, unitarity, and square-integrability to be established a priori, so that the $\mathscr{R}^S(\mathbf{C}_\mu)$ provide a construction of the discrete series independent of Harish-Chandra's character theory (see [**Wall2**] for a careful account). That the $\mathscr{R}^S(\mathbf{C}_\mu)$ yield *all* discrete series is still beyond primarily algebraic methods. However, they do, and many other representations besides. For example, in [**VoZu**], all unitary representations with nonvanishing relative $(\mathbf{g}, \mathbf{k})$-cohomology are classified. These are the representations whose multiplicities in $L^2(\Gamma\backslash G)$ determine the Betti numbers of locally symmetric spaces, via Matsushima's formula [**BoWa, HoWa, Mats**].

4. Other directions and applications. The "applications" of Lie theory are diverse and many; but often it is absurd to speak of "applications" when the role of Lie theory is so basic and pervasive: although $\mathbf{R}^n$ is a Lie group, usually it is used in such a low-tech way that its Lie-theoretic properties are superfluous. But when we combine it with its character group to form the Heisenberg group acting on $L^2(\mathbf{R}^n)$, then its identity as a Lie group becomes more relevant. Similarly with linear algebra: it would be egregious to claim it completely as part of Lie theory, but as I hope was demonstrated in §1, the border beyond which one should definitely consider oneself in Lie-theoretic territory is easy to cross and not so far from the public entrance. There is a similar identity problem on the high-end: to what extent should one include the more exotic algebraic structures dreamed up in physics—Jordan algebras, Kac-Moody algebras, Lie "superalgebras", quantum groups, vertex algebras—in "Lie theory"? Thus a comprehensive survey of "applications of Lie theory" is not simply impossible, it is fruitless. Below I offer instead an eclectic set of examples that I hope hit a few of the high points. For some original and stimulating discussions of many applications of representation theory, not exclusively of Lie groups, see various books and articles of Mackey [**Mack1–3**].

4.1. *Combinatorics.* Representation theory, even just the finite-dimensional theory of the general linear group, is rife with combinatorial quantities. We illustrate with a few examples.

4.1.1. S-FUNCTIONS. Much of the combinatorics of symmetric functions, developed in the nineteenth century, found natural interpretations in terms of representations of GL_n when that subject began to be understood through the work of Schur [**Schu**] around the turn of the century. The symmetric functions known as S-functions or Schur functions [**Litt, Macd1**] were introduced by Jacobi, but have been named after Schur because of his interpretation of them as the characters of the irreducible "polynomial" representations of GL_n or U_n. There is a famous identity due to Cauchy [**Macd1, Weyl2**] that, when interpreted in terms of representation theory, yields one of the most useful formulations of the fundamental theorems of classical invariant theory.

We will give a brief explanation of Cauchy's identity and its representation-theoretic interpretation. For an extensive discussion of symmetric functions with applications to representation theory, I recommend [**Macd1**]. In this discussion, we will follow the notation of [**Macd1**], although it differs from our earlier notation.