$\exp tX$, $X \in \mathbf{p}$. It behaves very simply: it is essentially a gradient flow, with nondegenerate, isolated fixed points, parametrized by the Weyl group. In fact, the Bruhat decomposition arises as the Morse decomposition **[Miln2]** associated to this flow. Similar facts can be shown to hold for the holomorphic action of $\mathbf{C}^{\times}$ on projective varieties **[Carr]**. (Of course $B^{0+}\backslash \mathrm{SL}_n(\mathbf{R})$ is not a complex variety, but if we had been working with $\mathrm{SL}_n(\mathbf{C})$ rather than $\mathrm{SL}_n(\mathbf{R})$, we would have been dealing with the complex flag manifold, which is a complex variety.) The geometric analysis allows one to describe quite precisely the asymptotics of Toda trajectories.

(e) The Toda lattice can be solved quite explicitly **[Kost6, Syme1, GoWa]**. It turns out to be closely related to the famous "QR algorithm" for diagonalizing matrices. It is essentially a continuous-time version of this algorithm **[Syme2, DeNT, GoWa]**.

(f) It should be noted that Proposition 4.3.45 is a generalization of the results of the explicit calculations (4.3.10) through (4.3.21).

**Appendix 1: Basic concepts of representation theory.** An account of representation theory must begin with some basic definitions and some remarks about certain technical issues. These latter can be somewhat off-putting, but if openly acknowledged, their negative effects can be minimized. The reader is advised to skim this section, and refer to it as necessary. For greater detail on this basic material, we refer to **[FeDo, Gaal, Kiri, Lang1]**, etc.

A.1.1. Let $G$ be a Lie group. A *representation* $\rho$ of $G$ on a vector space $V$ is a homomorphism of $G$ into the group of invertible linear transformations of $V$:

$$\rho : G \to \mathrm{GL}(V). \tag{A.1.1.1}$$

To be complete, in referring to a representation, we should specify both $\rho$ and $V$, but often we will only specify $\rho$, letting $V$ be understood implicitly; or we may just specify $V$ and let $\rho$ be implicit, in which case we call $V$ a *G-module.*

A.1.2. Very often $V$ will be infinite dimensional, and then usually it is equipped with a topology. Although the case of greatest general interest is when $V$ is a Hilbert space, sometimes it is a Banach space, and it is not really possible to avoid considering situations when $V$ is only locally convex. Whatever the topology of $V$, one wants to put a continuity condition on $\rho$. The correct one is that $\rho$ should be *strongly continuous*:

(A.1.2.1) The map $g \to \rho(g)v$, from $G$ to $V$, should be continuous for all $v \in V$.

REMARK. If $V$ is a Banach space, one might be tempted to think the map $g \to \rho(g)$ should be continuous with respect to the norm topology on the operators on $V$. But this condition is far too restrictive, and hardly ever holds when $V$ is infinite dimensional.

*From now on, all representations under discussion will be understood to be strongly continuous, unless the contrary is specifically stated.*

A.1.3. A very important special class of representations are those for which $V$ is a Hilbert space, and $\rho(G)$ consists of unitary operators. These are called *unitary* representations.

A.1.4. Let $\rho$ be a representation of $G$ on the space $V$, and $\sigma$ another representation on the space $U$. Then a (continuous) linear map $T : V \to U$ is called an *intertwining operator* (or if we want to be modern, morphism of $G$-modules or *G-morphism*) if

$$\text{(A.1.4.1)} \qquad \sigma(g)T = T\rho(g), \qquad g \in G.$$

That is, if the diagram

$$\begin{array}{ccc} V & \xrightarrow{\rho(g)} & V \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ U & \xrightarrow{\sigma(g)} & U \end{array}$$

commutes.

A.1.5. With notation as in §A.1.4, if $T$ is a topological isomorphism between $V$ and $U$, we call $T$ an *equivalence*. If such a $T$ exists, we say $\rho$ and $\sigma$ are equivalent. Very often, we are interested in representations only up to equivalence, and often when we say "representation" we actually mean equivalence class of representations. Context should usually make this clear.

A.1.5.1. If $\rho$, $\sigma$ are unitary representations, then we feel best if an equivalence $T$ is also unitary, in which case we speak of *unitary equivalence*. An easy argument using the polar decomposition (cf. **[Gaal, Lang2, RiNa]**, etc.) of an operator guarantees that if $\rho$, $\sigma$ are unitary and equivalent, then they are unitarily equivalent.

A.1.5.2. The notion of equivalence is a vexed one, owing to the extraordinary variety of linear topological vector spaces one has available. For example, consider the unit circle $\mathbf{T}$. There are billions and billions of function spaces on $\mathbf{T}$, perhaps most prominently the spaces $L^p(\mathbf{T})$, on which $\mathbf{T}$ acts via its action on itself by translations. Clearly these spaces bear a strong family resemblance to each other. However, the representations of $\mathbf{T}$ on them are not equivalent in the sense just defined, and generations of abelian harmonic analysts have derived pleasure from sorting out the differences **[Katz]**. They still are.

However, at the more primitive level which characterizes much of nonabelian harmonic analysis, one often would like to ignore differences such as those between $L^p$ and $L^q$. For this purpose, various looser notions of equivalence have been formulated. None so far are terribly satisfactory for a wide range of applications, although there has been some progress made in particular areas, **[Fell1, 2, Warn]**. Currently this is an unsettled, highly technical topic, which we will mostly ignore. An exception is the notion of infinitesimal equivalence of Harish-Chandra modules, which we will explain in §A.1.20 (see also §3.6.5).

A.1.6. Let $\rho$ be a representation of $G$ on $V$. Let $V_1 \subseteq V$ be a closed subspace. We say $V_1$ is *G-invariant* if $\rho(g)V_1 \subseteq V_1$ for all $g \in G$. If $V_1$ is $G$-invariant, then the map

$$\rho_1 : G \to \mathrm{GL}(V_1), \qquad \rho_1(g) = \rho(g) \mid V_1,$$

defined by restriction to $V_1$ is a representation of $G$ on $V_1$. We call $\rho_1$ arising in this way a *subrepresentation* of $\rho$.

A.1.7. Notations as in §A.1.6. If the only $G$-invariant subspaces of $V$ are $\{0\}$ and $V$, then we say $V$ is *irreducible.*

A.1.7.1. Irreducibility, like equivalence, is not a completely satisfactory notion. A better one for many purposes is *topological complete irreducibility* (t.c.i). We say $\rho$ is t.c.i. if for any finite-dimensional subspace $Y \subseteq V$, any linear map $T : Y \to V$, and any open neighborhood $U$ of $T$ in $\mathrm{Hom}(Y, V)$, we can find a linear combination $A = \sum a_i \rho(g_i)$ of elements of $\rho(G)$ such that $A_{|Y}$ is in $U$. That is, we can approximate an arbitrary operator of $\mathrm{End}(V)$ arbitrarily closely on finite-dimensional subspaces of $V$ by linear combinations from $\rho(G)$.

A.1.7.2. The condition t.c.i. implies irreducibility trivially. For unitary representations, irreducibility implies t.c.i. **[Warn]**.

A.1.7.3. The set of unitary equivalence classes of irreducible unitary representations of a group $G$ is called the *unitary dual* of $G$, and denoted $\widehat{G}$.

A.1.7.4. It is easy to check that if $\rho$ is a t.c.i. representation of $G$ on a space $V$, then the only operators on $V$ which commute with $\rho(G)$ are scalar multiples of the identity. In particular, if $ZG$ is the center of $G$, then $\rho(z)$, $z \in ZG$, must be a scalar multiple of the identity operator:

$$\rho(z) = \psi(z)I$$

for an appropriate complex number $\psi(z) \in \mathbf{C}^x$. It is immediate that $\psi : ZG \to \mathbf{C}^x$ must be a group homomorphism, often called a *quasicharacter.* We call $\psi$ the *central character* of $\rho$.

A.1.8. Notations as in §A.1.6. Let $V_2$ be another $G$-invariant subspace of $V$. Let $\rho_2$ be the subrepresentation of $V$ defined by $V_2$. Suppose that $V$ is the direct sum of $V_1$ and $V_2$, i.e., the natural map

$$\alpha : V_1 \oplus V_2 \to V,$$
$$\alpha(v_1, v_2) = v_1 + v_2, \qquad v_i \in V_i,$$

is a linear isomorphism. Then we say $\rho$ is the *direct sum* of the representations $\rho_1$ and $\rho_2$.

A.1.8.1. If $\rho$ is unitary, so that $V$ is a Hilbert space, let $V_1^\perp$ be the orthogonal complement of $V_1$. Then $V_1^\perp$ is also $G$-invariant. Hence, if a unitary representation is reducible, it decomposes as a direct sum.

A.1.9. Let $f$ be a complex-valued function on $G$. For $g \in G$, define the *left translation* of $f$ by $g$, $L_g(f)$, and the *right translation* of $f$ by $g$, $R_g(f)$, by the formulas

$$(\text{A.1.9.1}) \qquad L_g(f)(h) = f(g^{-1}h), \quad R_g(f)(h) = f(hg), \qquad g, h \in G.$$

(Actually, there is no need to require $f$ to be complex-valued for these formulas—it could be vector-valued, or even set-valued.) Let $\mathbf{C}^G$ be the vector space of all complex-valued functions on $G$ (no topology!). Then $L : g \to L_g$ and $R : g \to R_g$ are homomorphisms of $G$ into $\mathrm{GL}(\mathbf{C}^G)$. They are called the *left-regular* and *right-regular* representations of $G$.

Suppose $Y \subseteq \mathbf{C}^G$ is some space of functions on $G$, invariant by left translations by $G$, and equipped with a topology such that the restriction of $L$ to $Y$ is strongly continuous (cf. (A.1.2.1)). As a condition of nondegeneracy, to ensure that $Y$ consists of "most" functions on $G$, we will require that $Y$ contain $C_c^\infty(G)$, the space of smooth functions of compact support. We will refer to the restriction of $L$ to $Y$ as the *$Y$-left-regular representation*, or the *left-regular representation on $Y$*. Similar definitions apply to the right-regular representation. Frequently considered examples are $C_c^\infty(G)$ itself: $C_c(G)$, the continuous functions of compact support; $C_0(G)$, the continuous functions vanishing at $\infty$; and $L^p(G)$, $1 \leq p < \infty$ (these $L^p$ spaces are understood to be with respect to the left-invariant Haar measure on $G$ **[HeRo, Loom, Nach]**. Note that the action of $G$ on $L^\infty(G)$, and if $G$ is noncompact, even on $\mathbf{C}(G)$, the bounded continuous functions, is *not* strongly continuous.

A.1.10. Given a representation $\rho$ of $G$ on a space $V$, we can define an action $\rho^*$ of $G$ on the dual space $V^*$ of $V$ by the formula

$$(\text{A.1.10.1}) \qquad \rho^*(g)(\lambda)(v) = \lambda(\rho(g^{-1})v), \qquad \lambda \in V^*, v \in V, g \in G.$$

The action $\rho^*$ may not be strongly continuous. For example, if $\rho$ is the left-regular representation on $L^1(G) = V$, then $V^* = L^\infty(G)$, and the resulting $\rho^*$, which is just the left-regular representation on $L^\infty(G)$, is not strongly continuous. However, if $V$ is, for example, a reflexive Fréchet space, then $\rho^*$ will be strongly continuous **[Moor3, Warn]**. Whenever it is, we call $\rho^*$ the *contragredient* representation to $\rho$.

A.1.11. Let $\rho$ be a representation of $G$ on $V$. Select $v \in V$ and $\lambda \in V^*$.

The function on $G$ defined by

$$\varphi_{u,\lambda}(g) = \lambda(\rho(g)(u)) \tag{A.1.11.1}$$

is called the *matrix coefficient* of $\rho$ defined by $\lambda$ and $u$. It is a continuous function on $G$. If $\rho$ is unitary, or more generally if $\rho$ is a bounded representation (i.e., $|||\rho(g)||| \leq M$ for all $g \in G$ and some number $M$) on a Banach space, then $\varphi_{u,\lambda}$ is a bounded function on $G$.

A trivial formal calculation shows that

$$L_g\varphi_{u,\lambda} = \varphi_{u,\rho^*(g)(\lambda)}, \qquad R_g\varphi_{u,\lambda} = \varphi_{\rho(g)(u),\lambda}. \tag{A.1.11.2}$$

Thus the maps

$$\Phi_\lambda : u \to \varphi_{u,\lambda}, \quad \Phi_u^* : \lambda \to \varphi_{u,\lambda}, \qquad u \in V, \lambda \in V^*,$$

are intertwining maps between $\rho$ and the right-regular representation, and between $\rho^*$ and the left-regular representation, respectively. (We will be vague here about exactly which space of functions we are using for our regular representations.) Thus matrix coefficients serve as a bridge between the abstract world of representations and the concrete world of functions, specifically on $G$. In doing this, they play a pivotal role in representation theory. They also provide an intimate link with classical mathematics, as most of the special functions of the nineteenth century physics are matrix coefficients of appropriate representations of appropriate groups [**Mill1**, **Vile**].

A.1.12. If $\rho$ is a representation of $G$ on $V$, then the linear span of $\rho(G)$ is an algebra, as is its closure. It is often useful to be able to produce fairly general elements in this algebra. This is done by constructing the "integrated form" of $\rho$.

Let $dg$ be the left-invariant Haar measure on $G$ [**HeRo**, **Loom**, **Nach**]. For $f_1, f_2$ in $C_c^\infty(G)$, we define the convolution

$$f_1 * f_2(h) = \int_G f_1(g) f_2(g^{-1}h)\, dg = \int_G f_1(g) L_g(f_2)(h)\, dg. \tag{A.1.12.1}$$

This product turns $C_c^\infty(G)$ into an associative algebra. There is also an involutive, conjugate-linear, antiautomorphism (in brief: an involution) on $C_c^\infty(G)$ defined (in the case when $G$ is unimodular [**HeRo**, **Loom**, **Nach**]) by the formula

$$f^*(g) = \overline{f(g^{-1})}. \tag{A.1.12.2}$$

Here the overbar indicates complex conjugation. Note that

$$\check{f}(g) = f(g^{-1}) \tag{A.1.12.3}$$

is a complex linear antiautomorphism of $G$.

Given a representation $\rho$ of $G$ on a space $V$, and a function $f$ in $C_c^\infty(G)$, define $\rho(f)$ by the recipe

$$\rho(f)(v) = \int_G f(g)\rho(g)(v)\, dg, \qquad v \in V. \tag{A.1.12.4}$$

For this integral to be defined, the space $V$ must have some mild completeness properties [**Moor3**]. It is more than enough that $V$ be Fréchet. One checks by a formal computation that this definition of $\rho(f)$ defines an algebra homomorphism from $C_c^\infty(G)$, with convolution as product, to the algebra $\mathrm{End}(V)$ of continuous linear transformations on $V$.

*Unless the contrary is stated, we will understand that* $\rho$ *has an integrated form defined by* (A.1.12.4).

If $\rho$ is unitary, then formula (A.1.12.4) defines a *-homomorphism of $C_c^\infty(G)$ into $\mathrm{End}(V)$:

$$\rho(f^*) = \rho(f)^*, \tag{A.1.12.5}$$

where the $^*$ on the right-hand side of the equation indicates the adjoint of an operator on a Hilbert space (and the $^*$ on the left is as in formula (A.1.12.2)).

The algebra $C_c^\infty(G)$ can be completed in various norms to produce a Banach $^*$-algebra. The two most commonly considered ones are

$$\begin{aligned} &\text{The } L^1\text{-norm: } \|f\|_1 = \int_G |f(g)|\,dg, \\ &\text{The } C^*\text{-norm: } \|f\|_* = \sup\{|||\rho(f)||| : \rho \text{ unitary}\}. \end{aligned} \tag{A.1.12.6}$$

An easy estimate shows that if $\rho$ is a representation of $G$ by isometries on a Banach space, then $\rho$ is norm decreasing with respect to $\| \ \|_1$:

$$|||\rho(f)||| \le \|f\|_1 \quad \text{if } \rho \text{ is isometric in a Banach space.} \tag{A.1.12.7}$$

Here $||| \ |||$ indicates the norm of an operator on the space of the representation. Estimate (A.1.12.7) shows in particular that the supremum involved in defining $\| \ \|_*$ exists, such that

$$\|f\|_* \le \|f\|_1. \tag{A.1.12.8}$$

The completion of $C_c^\infty(G)$ with respect to $\| \ \|_*$ is a $C^*$-algebra [**FeDo, Gaal**].

The algebra $C_c^\infty(G)$ does not have an identity element. However, it does have an *approximate identity* or *Dirac sequence* (cf. [**HeRo, Lang1, 2**], etc.). This is a sequence of functions $\{f_n\}$ such that

$$\lim_{n\to\infty} f_n * \varphi = \varphi, \qquad \varphi \in C_c^\infty(G). \tag{A.1.12.9}$$

One constructs a Dirac sequence by choosing functions $f_n$ whose support is concentrated in smaller and smaller neighborhoods of the identity. Such a sequence is of course far from unique. When convenient, one can assume that $f_n$ is nonnegative, and that $\int_G f_n(g)\,dg = 1$ for all $n$. Given an approximate identity $\{f_n\}$, and a representation $\rho$ of $G$, one has

$$\lim_{n\to\infty} \rho(f_n)(v) = v. \tag{A.1.12.10}$$

In particular, the span of vectors of the form $\rho(f)v$, for $f \in C_c^\infty(G)$ and $v \in V$, is dense in $V$.

A.1.13. The construction of A.1.12, with $\mathbf{C}_c^\infty(G)$ replaced by $\mathbf{C}_c(G)$, works for a general locally compact group. For a Lie group $G$, one also has the Lie algebra $\mathrm{Lie}(G)$, and its universal enveloping algebra, and it is very much of the essence to represent them also. The procedure for doing this is somewhat more involved than in §A.1.12, since the operators of $\mathrm{Lie}(G)$ will typically be unbounded if $V$ is infinite dimensional.

Consider $X \in \mathrm{Lie}(G)$ and the associated one-parameter group $\{\exp tX\} \subseteq G$. Then $\rho(\exp tX)$ is a one-parameter group in $\mathrm{GL}(V)$, and $\rho(X)$ should be the infinitesimal generator of this one-parameter group. The procedure for finding $\rho(X)$ is clear from §2 (see formula (2.5.1)). We set

$$\rho(X)(v) = \lim_{t\to 0} \frac{\rho(\exp tX)v - v}{t} \tag{A.1.13.1}$$

for any $v$ for which this exists. (To be consistent with formula (2.5.1), we should denote $\rho(X)$ by $d\rho(X)$; but we now drop the $d$.) The formula

$$\rho(\exp sX)v - v = \int_0^s \rho(\exp tX)(\rho(X)v)\,dt \tag{A.1.13.2}$$

guarantees that $\rho(X)$ has a closed graph. If $f \in C_c^\infty(G)$, then one checks directly from the definitions that $\rho(X)\rho(f)v$ exists, for any $v \in V$; precisely, one has the formula

$$\rho(X)\rho(f)(v) = \rho(X(f))(v), \tag{A.1.13.3}$$

where $X(f)$ denotes the usual operation of differentiating the smooth function $f$ with respect to the vector field $X$. From the final result in §A.1.12, it follows that $\rho(X)$ is densely defined. Thus $\rho(X)$ is a closed, densely defined operator on $V$.

In fact, much more is true. The span of the vectors $\rho(f)(v)$, $f \in C_c^\infty(G)$, $v \in V$, form a dense subspace on which all $\rho(X)$, $X \in \mathrm{Lie}(G)$, are defined; and moreover, this subspace is stable under the $\rho(X)$'s by formula (A.1.13.3). This suggests the following construction. Define $V^\infty$ to be the subspace of $V$ consisting of vectors such that, for any sequence $X_1 X_2 \cdots X_n$ of elements in $\mathrm{Lie}(G)$, the composition $\rho(X_n)\rho(X_{n-1})\cdots\rho(X_1)(v)$ is defined. Topologize $V^\infty$ by requiring that a new $\{v_\alpha\}$ converges to a limit $v_0$ in $V^\infty$ if and only if, for all sequences of $X_i$'s, as above $\rho(X_n)\rho(X_{n-1})\cdots \rho(X_1)(v_\alpha)$ converges to $\rho(X_n)\cdots\rho(X_1)(v_0)$ in $V$. Then one shows

(A.1.13.4) Scholium: The space $V^\infty$ is dense in $V$; it is stable under $\rho(X)$, $X \in \mathrm{Lie}(G)$; also, it is stable under $G$; and the action of $G$ on $V^\infty$ is strongly continuous.

We call $V^\infty$ the space of *smooth vectors*, and we call the action of $G$ on $V^\infty$ the *smooth representation associated to* $\rho$. This is denoted sometimes by $\rho^\infty$, or sometimes simply again by $\rho$.

If $V = V^\infty$ (as topological vector spaces) we say $\rho$ is a *smooth representation*. It is clear that $(V^\infty)^\infty = V^\infty$, so that $\rho^\infty$ is always a smooth representation.

If $V$ is complete, then the fact that the $\rho(X)$ have closed graphs allows one to conclude that $V^\infty$ is complete. Since, in the definition of $V^\infty$ it would have sufficed to choose the elements $X_i$ from a basis of $\mathrm{Lie}(G)$, we see that if $V$ is Fréchet, in particular, if $V$ is Banach, then $V^\infty$ is Fréchet.

We have seen that vectors of the form $\rho(f)v$, $f \in C_c^\infty(G)$, $v \in V$, belong to $V^\infty$. The span of these vectors is known as the *Gårding space.* For a long time the relation between $V^\infty$ and the Gårding space was unclear. Then Dixmier and Malliavin showed they were equal **[DiMa]**.

REMARK. If $\mathbf{g}$ is a Lie algebra, the associative algebra generated by $\mathbf{g}$ subject to the relations $xy - yx = [x, y]$ (where the left-hand side indicates the usual commutator in an associative algebra, and the right-hand side is the Lie bracket in $\mathbf{g}$) is called the *universal enveloping algebra* of $\mathbf{g}$. We will denote it by $\mathscr{U}(\mathbf{g})$. Any representation

$$\rho : \mathbf{g} \to \mathrm{End}(V)$$

of $\mathbf{g}$ on a vector space $V$ automatically extends to a homomorphism

$$\rho : \mathscr{U}(\mathbf{g}) \to \mathrm{End}(V)$$

of associative algebras. Thus, in particular, in the context of this section, we have an action of $\mathscr{U}(\mathrm{Lie}(G))$ on $V^\infty$.

A.1.14. A major method for constructing representations is by *induction.* The study of induced representations for locally compact groups was pioneered by Mackey (cf. **[Mack4, FeDo, Warn]**, etc.).

Let $H \subseteq G$ be a closed subgroup. Let $\sigma$ be a representation of $H$ on a space $U$. Define $C_c^\infty(H\backslash G; \sigma)$ to be the space of smooth functions from $G$ to $U$ whose support lies in some set of the form $HS$, $S$ compact, and which transform by $\sigma$ under left translations by $H$:

$$\varphi(hg) = \sigma(h)\varphi(g), \qquad h \in H,\ g \in G,\ \varphi \in C_c^\infty(H\backslash G; \sigma). \tag{A.1.14.1}$$

It is easy to check that $C_c^\infty(H\backslash G; \sigma)$ is invariant under right translations by $G$. Further $C_c^\infty(H\backslash G; \sigma)$ comes equipped with a standard topology **[Warn]**, and it is not hard to see that with respect to this topology the action of $G$ is strongly continuous. The representation of $G$ so defined will be called the representation of $G$ *induced* by $\sigma$, or more particularly, the $C_c^\infty$*-induced representation* induced by $\sigma$. It will be denoted by $\mathrm{ind}_H^G \sigma$, or $C_c^\infty - \mathrm{ind}_H^G \sigma$ if we wish to be more specific. More generally, if $Y$ is a space obtained by completing $C_c^\infty(H\backslash G; \sigma)$ in some topology weaker than its natural topology, we will also call the action of $G$ on $Y$, extended from $C_c^\infty(H\backslash G; \sigma)$ by continuity, a representation induced by $\sigma$.

A.1.15. Continue the notations of §A.1.1.4. Consider the space $C^\infty(G; U)$ of smooth functions on $G$ with values in $U$. We can consider two actions of $H$ on $C^\infty(G; U)$, the action by left translations, as defined in (A.1.9.1), and the action by transforming the values of elements of $C^\infty(G; U)$ by $\sigma$.

These two actions commute with the action of $G$ by right translations and they also commute with each other, so we can form their "tensor product," which we will denote by $\sigma \otimes R$. Explicitly in formulas, these actions are given by

$$\begin{aligned} &L_h f(g) = f(h^{-1}g), \qquad f \in C^\infty(G; U), \\ &\sigma(h)(f)(g) = \sigma(h)(f(g)), \\ &(\sigma \otimes L)(h)(f)(g) = \sigma(h)(f(h^{-1}g)). \end{aligned} \tag{A.1.15.1}$$

In terms of these actions, we can see that $C_c^\infty(H\backslash G; \sigma)$ is defined by a condition of invariance under the action $\sigma \otimes L$ of $H$, as well as a support condition relative to $H\backslash G$.

Let $C_c^\infty(G; U)$ be the elements of $C^\infty(G; U)$ with compact support. For $\varphi \in C_c^\infty(G; U)$, define a function $p_\sigma(\varphi)$ in $C^\infty(G; U)$ by the recipe

$$p_\sigma(\varphi)(g) = \int_H \sigma(h)^{-1} \varphi(hg)\, d_r h, \tag{A.1.15.2a}$$

where $d_r h$ is a right-invariant Haar measure on $H$. In terms of notations (A.1.15.1), we may write

$$p_\sigma(\varphi) = \int_H \sigma \otimes L(h)(\varphi)\, d_l h, \tag{A.1.15.2b}$$

where $d_l h = d_r h^{-1}$ is a left-invariant Haar measure on $H$. Note that we can also write

$$d_l h = \delta_H(h)^{-1} d_r h, \tag{A.1.15.3}$$

where $\delta_H$ is the modular function of $H$ **[Gaal, Loom, Weil4]**. From formulas (A.1.15.2) and (A.1.15.3), it is a simple matter to check that

$$\sigma \otimes L(h)(p_\sigma(\varphi)) = p_\sigma(f), \; h \in H, \; \varphi \in C_c^\infty(G; U), \tag{A.1.15.4a}$$

$$p_\sigma(\sigma \otimes L(h)(\varphi)) = \delta_H(h) p_\sigma(\varphi). \tag{A.1.15.4b}$$

Since it is clear that $p_\sigma(\varphi)$ will have support contained in $H \operatorname{supp}(\varphi)$, we see that formula (A.1.15.4a) is just the statement that $p_\sigma(\varphi) \in C_c^\infty(H\backslash G; \sigma)$. We conclude that $p_\sigma$ defines a linear map

$$p_\sigma : C_c^\infty(G; U) \to C_c^\infty(H\backslash G; \sigma). \tag{A.1.15.5}$$

Both these spaces have standard topologies, and it is easy to check that $p_\sigma$ is a continuous map. The results of Dixmier and Malliavin **[DiMa]** imply that if $\sigma$ is smooth and $U$ is Fréchet, then the map $p_\sigma$ is surjective. In particular, if $\sigma$ is finite dimensional, then $p_\sigma$ is surjective (but this is elementary).

Regarding the kernel of $p_\sigma$, the formula (A.1.15.4b) makes it clear that functions of the form

$$(\sigma \otimes L)(h)(\varphi) - \delta_H(h)(\varphi), \qquad \varphi \in C_c^\infty(G; U),$$

are sent to zero by $p_\sigma$. If we let $h$ vary in a one-parameter group and differentiate at the origin, we may conclude that

$$(\sigma(x) + L(x) - \delta_H(x))\varphi \in \ker p_\sigma \tag{A.1.15.6}$$

for all $\varphi \in C_c^\infty(G; U)$ and $x \in \mathrm{Lie}(H)$. (Here we are following the notation of (A.1.13.1) for the action of $\mathrm{Lie}(H)$.)

LEMMA A.1.15.7. *If $\sigma$ is finite-dimensional, then every element of* $\ker p_\sigma$ *is a finite sum of functions of the form* (A.1.15.6).

This lemma is essentially a generalization of the Poincaré Lemma [**Gold**, **Ster**], for forms of top degree.

Let $\lambda$ be a linear functional on $C_c^\infty(H\backslash G; \sigma)$. Then $\lambda \circ p_\sigma$ is a linear functional on $C_c^\infty(G; U)$—a "$U^*$-valued distribution." Using formulas (A.1.15.4) and (A.1.10.1) we can compute that

$$\begin{aligned}(\sigma \otimes L)^*(h)((\lambda \circ p_\sigma)(\varphi)) &= (\lambda \circ p_\sigma)((\sigma \otimes L)(h)^{-1}\varphi)\\ &= \lambda(p_\sigma((\sigma \otimes L)(h)^{-1}(\varphi)))\\ &= \lambda(\delta_H(h)^{-1} p_\sigma(\varphi)) = \delta_H(h)^{-1}\lambda \circ p_\sigma(\varphi), \qquad h \in H,\ \varphi \in C_c^\infty(G; U).\end{aligned}$$

In other words, the distribution $\mu = \lambda \circ p_\sigma$ satisfies the transformation law

$$\delta_H(h)(\sigma \otimes L)^*(h)(\mu) = \mu. \tag{A.1.15.8}$$

Suppose on the other hand that we have a functional $\mu$ on $C_c^\infty(G; U)$ which satisfies the transformation law (A.1.15.8). By a differentiation, we conclude $\mu$ vanishes on functions of the form (A.1.15.6). Thus the following statement follows directly from Lemma A.1.15.7.

COROLLARY A.1.15.9. *Suppose $\sigma$ is a smooth representation on a Fréchet space $U$. Then any distribution $\mu$ on $C_c^\infty(G; U)$ satisfying the transformation law* (A.1.15.8) *is of the form $\mu = \lambda \circ p_\sigma$ for suitable $\lambda \in C_c^\infty(H\backslash G; \sigma)$.*

This corollary is the basic fact behind various Frobenius reciprocity statements (see [**Gaal**, **Knap2**, **Warn**], etc.).

As an example of Corollary A.1.15.9, consider right-invariant Haar measure $d_r g$ on $G$. Using formulas (A.1.15.3) and (A.1.10.1), we compute that

$$L^*(g_1)(d_r g) = \delta_G(g_1)^{-1} d_r g, \qquad g_1 \in G. \tag{A.1.15.10}$$

From Corollary A.1.15.9, we conclude that the measure $d_r g$ factors to the space $C_c^\infty(H\backslash G; \delta_H/\delta_G)$; in other words, there is a right $G$-invariant functional on this space. (Note that if $\delta_H = \delta_{G|H}$, this amounts to the statement that $H\backslash G$ carries a $G$-invariant measure [**Gaal**, **Weil4**, **FeDo**], etc.)

A.1.16. Isometric Banach space representations, especially unitary representations, are of particular interest, so we want to know how they fare with respect to induction. For this, it is convenient to deal with a class of representations slightly more general than isometric ones. Let us call a representation

$\rho$ of $G$ on a Banach space $V$ *quasi-isometric* if it satisfies the following mutually equivalent conditions

(A.1.16.1)
(a) $\rho(g)$ is a scalar multiple of an isometry of $V$ for all $g$ in $G$,
(b) $|||\sigma(g)|||\ |||\sigma(g)^{-1}||| = 1$. Here $|||\ |||$ denotes the operator norm in End $V$.
(c) $\sigma(g) = \alpha(g)\sigma_1(g)$, where $\sigma_1$ is an isometric representation of $G$ on $V$, and $\alpha\colon G \to \mathbf{R}^{+\times}$ is a continuous homomorphism, i.e., a real-valued quasicharacter on $G$.

Note that, in situation (A.1.16.1)(c), we have

(A.1.16.2) $$\alpha(g) = |||\sigma(g)|||.$$

Thus $|||\sigma(g)|||$ is a quasicharacter; we call it the *dilation character* of $\rho$.

If $\rho$ is quasi-isometric, and $\beta : G \to \mathbf{C}^{\times}$ is a quasicharacter of $G$, then $\beta \otimes \rho$, defined by

(A.1.16.3a) $$\beta \otimes \rho(g)(v) = \beta(g)\rho(g)(v),$$

is also quasi-isometric, with dilation character

(A.1.16.3b) $$|||\beta \otimes \rho||| = |\beta|\ |||\rho|||,$$

where $|\beta|$ denotes the absolute value of $\beta$.

Let $H \subseteq G$ be a closed subgroup, and let $\sigma$ be a quasi-isometric representation of $H$ on a Banach space $U$. Since $\sigma$ is quasi-isometric, we see that for any $f \in C_c^\infty(H\backslash G;\sigma)$, the function $\|f\|$, defined by

(A.1.16.4) $$\|f\|(g) = \|f(g)\|,$$

where $\|\ \|$ is the norm on $U$, will belong to $C_c(H\backslash G; \|\sigma\|)$. Suppose that $\|\sigma\| = (\delta_H/\delta_G)^{1/p}$ for some real number $p \geq 1$. Then $\|f\|^p \in C_c(H\backslash G;\delta_H/\delta_G)$. According to the final remark of §A.1.15, there is a right-invariant functional on $C_c(H\backslash G;\delta_H/\delta_G)$, a projection of right-invariant Haar measure on $G$. Let us denote it by $\varphi \to \int_{H\backslash G}\varphi\, d\dot{g}$. Then the recipe

(A.1.16.5) $$\|f\|_p = \left(\int_{H\backslash G} \|f\|^p\, d\dot{g}\right)^{1/p}$$

defines a $G$-invariant norm on $C_c^\infty(H\backslash G;\sigma)$, which may be completed to define an isometric representation of $G$.

The above construction involves a particular assumption on the dilation character $\|\sigma\|$ of the representation $\sigma$. However, we see that for any quasi-isometric representation $\sigma$ of $H$, the representation $((\delta_H/\delta_G)^{1/p}\|\sigma\|^{-1}) \otimes \sigma$ will satisfy the assumption. Thus we have

**Proposition** A.1.16.6. *If $\sigma$ is a quasi-isometric representation of the closed subgroup $H$ on the Banach space $U$, then formula* (A.1.16.5) *defines a G-invariant norm on* $C_c^\infty(H\backslash G; ((\delta_H/\delta_G)^{1/p}\|\sigma\|^{-1})\otimes\sigma)$.

The isometric representation of $G$ that results from completing the norm of Proposition A.1.16.6 will be denoted

$$p - \mathrm{ind}_H^G\ \sigma \tag{A.1.16.7}$$

and will be called the *p-normalized induced representation* derived from $\sigma$.

Finally, suppose that $\sigma$ is a unitary representation of $H$. Let $(\ ,\ )$ denote the $H$-invariant inner product which defines the norm on $U$. Then if $f_1, f_2$ are in $C_c^\infty(H\backslash G; (\delta_G/\delta_H)^{1/2}\otimes\sigma)$, we see that the inner product

$$(f_1, f_2)(g) = (f_1(g), f_2(g))$$

will be a scalar-valued function in $C_c^\infty(H\backslash G; \delta_G/\delta_H)$. Hence the recipe

$$(f_1, f_2)^\sim = \int_{H\backslash G}(f_1, f_2)\,d\dot{g} \tag{A.1.16.8}$$

defines a $G$-invariant inner product on $C_c^\infty(H\backslash G; (\delta_G/\delta_H)^{1/2}\otimes\sigma)$. Further, the norm defined by this inner product is clearly the norm attached to $2 - \mathrm{ind}_H^G\,\sigma$. In summary, we have shown

**Corollary** A.1.16.9. *If $\sigma$ is a unitary representation of $H$, then $2 - \mathrm{ind}_H^G\,\sigma$ is a unitary representation of $G$, with invariant inner product defined by* (A.1.16.8).

A.1.17. Let $V$ be a vector space and $V^*$ its dual. Given $v \in V$ and $\lambda \in V^*$, one can form the *dyad* $E_{v,\lambda}$, which is an operator on $V$, by the formula

$$E_{v,\lambda}(x) = \lambda(x)v, \qquad x \in V. \tag{A.1.17.1}$$

The bilinear map $(v, \lambda) \to E_{v,\lambda}$ extends to an embedding

$$E : V\otimes V^* \hookrightarrow \mathrm{End}(V) \tag{A.1.17.2}$$

of the algebraic tensor product into the algebra of linear transformations on $V$. The image $E(V\otimes V^*)$ consists of all operators of finite rank on $V$. If $V$ is a locally convex topological vector space and $V^*$ is the topological dual of continuous linear functionals on $V$, then the image of $E$ is the algebra of continuous finite rank operators on $V$.

On $V\otimes V^*$ there is defined a canonical linear functional induced by the canonical bilinear pairing between $V$ and $V^*$. When considered as a function of operators, this functional is called the *trace* and is denoted by tr. We have the formula

$$\mathrm{tr}\left(\sum E_{v_i,\lambda_i}\right) = \sum \lambda_i(v_i). \tag{A.1.17.3}$$

Suppose $V$ is a Banach space. We can define a norm $||| \; |||_1$ on $E(V \otimes V^*)$ by the rule

(A.1.17.4)
$$\|T\|_1 = \inf\left\{\sum_i \|v_i\| \, \|\lambda_i\| : \sum E_{v_i, \lambda_i} = T\right\}.$$

We call this the *trace norm* on $E(V \otimes V^*)$. It is easy to check that the trace norm dominates the usual operator norm on $E(V \otimes V^*)$. Hence Cauchy sequences with respect to $\| \; \|_1$ in $E(V \otimes V^*)$ will also be Cauchy with respect to $\| \; \|$, so the Banach space completion of $E(V \otimes V^*)$ can be regarded as a certain space of operators on $V$. We call it $\mathscr{T}(V)$, the space of *trace class* or *nuclear operators* on $V$ (cf. **[CoGr, Gaal, Lang1, 2]**, etc.). In fact, $\mathscr{T}(V)$ is a two-sided ideal in $\mathrm{End}(V)$, and one can easily check that

(A.1.17.5)
$$|||ATB|||_1 \leq |||A||| \; |||B||| \; |||T|||_1, \qquad A, B \in \mathrm{End}(V), \; T \in \mathscr{T}(V).$$

Here $||| \; |||$ denotes the usual operator norm on $\mathrm{End}(V)$.

Since the trace linear functional on $E(V \otimes V^*)$ is clearly dominated by the trace norm, it extends by continuity to a linear function, still called the trace, on $\mathscr{T}(V)$.

A.1.18. Let $\rho$ be a representation of $G$ on a Banach space $V$. It may happen that, for some element $X$ in the universal enveloping algebra $\mathscr{U}(\mathrm{Lie}(G))$ (see the Remark at the end of §A.1.13) the operator $\rho(X)^{-1}$ is trace class. By this we mean

(A.1.18.1)
(i) As an operator on $V^\infty$, $\rho(X)$ is invertible.
(ii) The inverse operator $\rho(X)^{-1}$ extends to a continuous operator on $V$.
(iii) This continuous extension is in the space $\mathscr{T}(V)$ of trace class operators on $V$.

If this happens, we call $\rho$ a *strongly trace class representation.* Many Lie groups, including all nilpotent groups and all semisimple groups (with a finite number of connected components), have all their irreducible representations strongly trace class **[CoGr, Warn]**.

Under the conditions of the previous paragraph, consider $f \in C_c^\infty(G)$. We can write

$$\rho(f) = \rho(X)^{-1}\rho(X)\rho(f) = \rho(X)^{-1}\rho(L(X)(f)).$$

We conclude that $\rho(f)$ is trace class; further, $f \to \rho(f)$ is continuous from $C_c^\infty(G)$ to $\mathscr{T}(V)$. In particular, the functional

(A.1.18.2)
$$\theta_\rho(f) = \mathrm{tr}\,\rho(f)$$

is a distribution on $G$. We call $\theta_\rho$ the *distributional character*, or simply the character, of $\rho$.

A.1.19. Let $K$ be a compact group, and let $\rho$ be a representation of $K$ on a vector space $V$. A vector $v$ in $V$ is called *K-finite* if the span of $\{\rho(k)v, k \in K\}$, the $K$-transforms of $v$, is finite dimensional. The set of all $K$-finite vectors in $V$ is a subspace of $V$, denoted $V_K$. An argument using approximate identities and the Peter-Weyl Theorem shows that $V_K$ is dense in $V$ (cf. [**Lang1**, **Warn**, **Knap2**], etc.).

Let $\sigma \in \widehat{K}$ be an (isomorphism class of) irreducible representation(s) of $K$. A vector $v$ in $V$ is of *type* $\sigma$ if the span of the $K$-transforms is an irreducible representation isomorphic to $\sigma$. The *$\sigma$-isotypic component* of $V$ is the span of all vectors of type $\sigma$. It is denoted $V_\sigma$. The Peter-Weyl Theorem (cf. Theorem 3.5.4.23) provides a function $e_\sigma$ such that $\rho(e_\sigma)$ is a projection from $V$ to $V_\sigma$. It follows that *every* vector in $V_\sigma$ is $K$-finite, and any finite-dimensional, $K$-invariant subspace of $V_\sigma$ is isomorphic to a direct sum of copies of $\sigma$. The ratio (perhaps infinite) $(\dim V_\sigma)/\dim \sigma$ is called the *multiplicity* of $\sigma$ in $V$.

We have

$$V_K = \sum_{\sigma \in \widehat{K}} V_\sigma.$$

This may be considered simply as the algebraic direct sum of the spaces $V_\sigma$, or, more elaborately, may be considered as a topological vector space which is the inductive limit over $\sum_{\sigma \in F} V_\sigma$, for finite sets $F \subseteq \widehat{K}$. Each $\sum_{\sigma \in F} V_\sigma$ is given its topology as a subspace of $V$.

A.1.20. Let $K$ be a compact subgroup of the Lie group $G$. Let $\rho$ be a representation of $G$ on a vector space $V$. Let $V_K$ be the subspace of $K$-finite vectors (cf. §A.1.19). Since $\mathrm{Lie}(G)$ is a finite-dimensional $K$-module under the adjoint action, and the action of $\mathrm{Lie}(G)$ on $V^\infty$ (cf. §A.1.13) can be expressed in terms of a map

$$\mathrm{Lie}(G) \otimes V^\infty \to V^\infty, \qquad x \otimes v \to \rho(x)v,$$

one can check that $V_K^\infty$ is invariant under $\mathscr{U}(\mathrm{Lie}(G))$, the universal enveloping algebra (cf. §A.1.13). Thus $V_K^\infty$ is a module for $K$ and for $\mathscr{U}(\mathrm{Lie}(G))$.

Let $\rho$, $\rho'$ be representations of $G$ on spaces $V$, $V'$. We say $V$, $V'$ are *infinitesimally equivalent* if there is a linear isomorphism

$$T : V_K^\infty \to V_K'^\infty$$

which intertwines the actions of $K$ and of $\mathscr{U}(\mathrm{Lie}(G))$ on these two spaces. The situation in which this notion is most often used is when $G$ is semisimple, $K$ is a maximal compact subgroup, and $V$, $V'$ are irreducible (t.c.i. Banach) representations of $G$. In these circumstances, the $K$-isotypic components $V_\sigma$ and $V'_\sigma$, $\sigma \in \widehat{K}$, are finite dimensional, by an early result of Harish-Chandra (cf. [**HaCh3**, **Gode2**, **Warn**, **Knap2**], etc.). In this case $V_K^\infty = V_K$ constitutes an algebraic skeleton around which $V$ is built, by means of completion with respect to some topology. Infinitesimal equivalence throws away the fuzz introduced with the topology, and considers only the algebraic core. At the current stage of semisimple harmonic analysis, which is still primarily concerned with individual irreducible representations, this is a very useful notion of equivalence.

**Appendix 2: Structure of real semisimple Lie algebras and Lie groups.**

A.2.1. *Cartan involution and invariant form.* We follow **[Knap2]** and **[Wall2]** by defining a reductive Lie group to be a closed subgroup $G$ of $\mathrm{GL}_n(\mathbf{R})$, for some $n$, which is left-invariant (as a set) by the "Cartan involution"

(A.2.1.1) $$\theta : g \to (g^t)^{-1}.$$

Here $g^t$ indicates the transpose of $g \in \mathrm{GL}_n(\mathbf{R})$. This definition allows one to short-circuit a lot of preliminary material and get to the essential facts fairly quickly.

The "infinitesimal automorphism" of $\mathbf{gl}_n(\mathbf{R}) \simeq M_n(\mathbf{R})$ corrresponding to $\theta$ as defined in (A.2.1.1) is

(A.2.1.2) $$\theta(x) = -x^t;$$

also known as the Cartan involution. Since it is of order two, its eigenvalues are simply $\pm 1$. The space of matrices fixed by $\theta$ is exactly the space of skew-symmetric matrices; this is also the Lie algebra $\mathbf{o}_n$ of the orthogonal group, the isometry group of the standard inner product on $\mathbf{R}^n$:

(A.2.1.3) $$\mathbf{o}_n = \{x \in M_n(\mathbf{R}) : x + x^t = 0\} = \{x \in M_n(\mathbf{R}) : \theta(x) = x\}.$$

The $-1$ eigenspace of $\theta$ is the space $\mathbf{s}$ of symmetric matrices

(A.2.1.4) $$\mathbf{s} = \{y \in M_n(\mathbf{R}) : y = y^t\}.$$

We have the direct sum decomposition

(A.2.1.5) $$\mathbf{gl}_n(\mathbf{R}) \simeq \mathbf{o}_n \oplus \mathbf{s}.$$

The summand $\mathbf{o}_n$ is a Lie algebra, the Lie algebra of the compact group $\mathrm{O}_n \subseteq \mathrm{GL}_n(\mathbf{R})$, but the summand $\mathbf{s}$ is not at all a Lie algebra: its commutators belong to $\mathbf{o}_n$. However, it is invariant under commutators from $\mathbf{o}_n$. In sum, we have the relations

(A.2.1.6) $$[\mathbf{o}_n, \mathbf{o}_n] \subseteq \mathbf{o}_n, \qquad [\mathbf{o}_n, \mathbf{s}] \subseteq \mathbf{s}, \qquad [\mathbf{s}, \mathbf{s}] \subseteq \mathbf{o}_n.$$

Here $[\mathbf{o}_n, \mathbf{o}_n]$ is interpreted as the linear span of commutators $[x, y]$, $x, y \in \mathbf{o}_n$, and similarly for the other expressions.

We consider on $\mathbf{gl}_n(\mathbf{R})$ the bilinear form

(A.2.1.7) $$B(x, y) = \mathrm{tr}(xy), \qquad x, y \in M_n(\mathbf{R}),$$

where tr indicates the trace function on matrices. We observe that $B(x, y)$ is invariant under conjugation:

(A.2.1.8) $$\begin{aligned} B(\mathrm{Ad}\, g(x), \mathrm{Ad}\, g(y)) &= \mathrm{tr}((gxg^{-1})(gyg^{-1})) \\ &= \mathrm{tr}(gxyg^{-1}) = \mathrm{tr}(xy) = B(x, y). \end{aligned}$$

The infinitesimal version of this is

$$B(\operatorname{ad} x(y), z) + B(x, \operatorname{ad} y(z)) = 0. \tag{A.2.1.9}$$

We observe that the summands $\mathbf{o}_n$ and $\mathbf{s}$ of decomposition (A.2.1.5) are orthogonal with respect to $B$. Furthermore, it is easy to check that $B$ is negative definite on $\mathbf{o}_n$, and positive definite on $\mathbf{s}$.

Now consider a Lie algebra $\mathbf{g} \subseteq \mathbf{gl}_n(\mathbf{R})$. Assume that $\mathbf{g}$ is invariant under the Cartan involution $\theta$, given in equation (A.2.1.2). We will refer to $\theta_{|\mathbf{g}}$ as the Cartan involution for $\mathbf{g}$.

Since $\mathbf{g}$ is invariant under $\theta$, we see that if we set

$$\mathbf{k} = \mathbf{o}_n \cap \mathbf{g}, \qquad \mathbf{p} = \mathbf{s} \cap \mathbf{g}$$

then we will have the direct sum decomposition

$$\mathbf{g} = \mathbf{k} \oplus \mathbf{p}. \tag{A.2.1.10}$$

The analogs of relations (A.2.1.6) will clearly hold for $\mathbf{k}$ and $\mathbf{p}$. In particular, $\mathbf{k}$ is a Lie subalgebra of $\mathbf{g}$, and $\mathbf{p}$ is a $\mathbf{k}$-module via the adjoint action. Also, $\mathbf{k}$ and $\mathbf{p}$ are orthogonal with respect to the restriction of the inner product $B$ of formula (A.2.1.7) to $\mathbf{g}$, and $B_{|\mathbf{k}}$ is negative definite while $B_{|\mathbf{p}}$ is positive definite. It follows that $B_{|\mathbf{g}}$ is nondegenerate. We may also conclude that if $K = \exp \mathbf{k}$ is the connected subgroup of $\mathrm{GL}_n(\mathbf{R})$ with Lie algebra $\mathbf{k}$, then $K$ has compact closure (since it will be contained in $O_n$). Furthermore, $\mathbf{k}$ is a maximal subalgebra of $\mathbf{g}$ with the property that it generates a bounded group; for if $\exp tx$, $x \in \mathbf{gl}_n(\mathbf{R})$, is a bounded one-parameter subgroup, the eigenvalues of $\exp tx$ must be of absolute value $1$, hence the eigenvalues of $x$ must be pure imaginary, hence $B(x, x) \leq 0$. However, any subspace of $\mathbf{g}$ strictly containing $\mathbf{k}$ must have nontrivial intersection with $\mathbf{p}$, on which $B$ is positive definite.

A.2.2. *Split Cartan subalgebras and restricted roots.* Elements of $\mathbf{p}$, being symmetric matrices, are diagonalizable with real eigenvalues, hence the same is true for their action on $\mathbf{g}$ by ad. For $x \in \mathbf{p}$, let $\mathbf{c}_{\mathbf{g}}(x) = \ker \operatorname{ad} x_{|\mathbf{g}}$ be the centralizer of $x$ in $\mathbf{g}$. It is easy to check that $\mathbf{c}_{\mathbf{g}}(x)$ is a Lie subalgebra of $\mathbf{g}$, and is stable under the Cartan involution. Thus $\mathbf{c}_{\mathbf{g}}(x) = (\mathbf{c}_{\mathbf{g}}(x) \cap \mathbf{k}) \oplus (\mathbf{c}_{\mathbf{g}}(x) \cap \mathbf{p})$. Clearly, $x \in \mathbf{c}_{\mathbf{g}}(x) \cap \mathbf{p}$, and is central in $\mathbf{c}_{\mathbf{g}}(x)$. If $\mathbf{c}_{\mathbf{g}}(x) \cap \mathbf{p}$ is not central in $\mathbf{c}_{\mathbf{g}}(x)$, we can choose another element, $x_2$, in $\mathbf{c}_{\mathbf{g}}(x)$ and look at its centalizer. Continuing in this way, we will arrive at a subalgebra

$$\mathbf{m} \oplus \mathbf{a} \tag{A.2.2.1}$$

such that

(i) $\mathbf{m} \subseteq \mathbf{k}$, $\mathbf{a} \subseteq \mathbf{p}$,
(ii) $\mathbf{m} \oplus \mathbf{a}$ is the centralizer of $\mathbf{a}$ in $\mathbf{g}$.

Because of the analogy between this construction and the construction of Cartan subalgebras for general Lie algebras, the abelian algebra $\mathbf{a}$ is called a *split Cartan subalgebra* of $\mathbf{g}$.

Let $\mathbf{a}$ be a split Cartan subalgebra of $\mathbf{g}$, as in (A.2.2.1), and consider the adjoint action of $\mathbf{a}$ on $\mathbf{g}$. Since elements of $\mathbf{a}$ are individually diagonalizable over $\mathbf{R}$, and since they commute, we can decompose $\mathbf{g}$ into simultaneous eigenspaces for $\operatorname{ad}\mathbf{a}$: If $x$ is a simultaneous eigenvector for all $\operatorname{ad} a$, then

$$\operatorname{ad} a(x) = \alpha(a)x,$$

where $\alpha(a)$ is the appropriate eigenvalue. Clearly $\alpha(a)$ depends linearly on $a$; that is, $\alpha$ belongs to $\mathbf{a}^*$, the dual of $\mathbf{a}$. Let $\Delta$ denote the set of nonzero elements of $\mathbf{a}^*$ which arise as the simultaneous eigenvalue function for some $x \in \mathbf{g}$. In parallel with the general situation, described in §2, the elements of $\Delta$ are called the *restricted roots*, or simply roots, of $\mathbf{g}$. We have a direct sum decomposition

(A.2.2.2) $$\mathbf{g} = \mathbf{g}_0 \oplus \sum_{\alpha\in\Delta} \mathbf{g}_\alpha,$$

where

(i) $\mathbf{g}_0 = \mathbf{m} \oplus \mathbf{a}$,
(ii) $\mathbf{g}_\alpha = \{x : \operatorname{ad} a(x) = \alpha(a)x, \text{ all } a \in \mathbf{a}\}, \alpha \in \Delta$.

The $\mathbf{g}_\alpha$ are called the *root spaces* for $\alpha$.

The Cartan involution $\theta$ normalizes $\mathbf{a}$, so it will preserve the decomposition (A.2.2.2). In fact, it is easy to check that, if $x \in \mathbf{g}_\alpha$ and $a \in \mathbf{a}$, then

$$\operatorname{ad} a(\theta(x)) = \theta(\operatorname{ad}\theta(a)(x)) = \theta(\alpha(-a)x) = -\alpha(a)\theta(x).$$

Thus

(A.2.2.3) $$\theta(\mathbf{g}_\alpha) = \mathbf{g}_{-\alpha}.$$

In particular,

(A.2.2.4) $$\mathbf{g}_\alpha \oplus \mathbf{g}_{-\alpha} = (\mathbf{g}_\alpha \oplus \mathbf{g}_{-\alpha}) \cap \mathbf{k} \oplus (\mathbf{g}_\alpha \oplus \mathbf{g}_{-\alpha}) \cap \mathbf{p}.$$

This decomposition allows us to show, by an argument analogous to the proof of uniqueness of Cartan subalgebras for complex Lie algebras, that the split Cartan subalgebra $\mathbf{a}$ of (A.2.2.1) is essentially unique.

LEMMA A.2.2.5. $\operatorname{Ad} K(\mathbf{a}) = \mathbf{p}$.

This can be proved by the following steps.

(i) Define an element $x \in \mathbf{p}$ to be *regular* if its centralizer in $\mathbf{p}$ is of minimal dimension, equivalently if it is abelian, equivalently if it is a split Cartan subalgebra. Denote the set of regular elements in $\mathbf{p}$ by $\mathbf{p}_{\text{reg}}$. It is open and dense in $\mathbf{p}$.

(ii) An easy computation of the derivative of $\operatorname{Ad} K$ at $x \in \mathbf{a} \cap \mathbf{p}_{\text{reg}}$, using equation (A.2.2.4) shows that $\operatorname{Ad} K(\mathbf{a} \cap \mathbf{p}_{\text{reg}})$ is open in $\mathbf{p}_{\text{reg}}$. Since the same reasoning applies to any putative other split Cartan subalgebra $\tilde{\mathbf{a}}$, it follows that $\operatorname{Ad} K(\mathbf{a} \cap \mathbf{p}_{\text{reg}})$ is also closed in $\mathbf{p}_{\text{reg}}$.

(iii) A computation similar to that of (ii), but at a nonregular element $x \in \mathbf{a}$, shows that $\mathbf{p} - \mathbf{p}_{\text{reg}}$ has codimension at least 2 in $\mathbf{p}$. Hence $\mathbf{p}_{\text{reg}}$ is connected, whence from (ii), $\text{Ad}\,K(\mathbf{a} \cap \mathbf{p}_{\text{reg}}) = \mathbf{p}_{\text{reg}}$. Taking closures gives the lemma.

Thus, in analogy with the decomposition (2.8.6) of a general complex Lie algebra with respect to a Cartan subalgebra, one finds the decomposition (A.2.2.2) of $\mathbf{g}$ into root spaces for $\mathbf{a}$ is essentially canonical (i.e., is unique up to conjugation). Further, one can find, in a similar fashion to §2.8, copies of $\mathbf{sl}_2$ (actually, $\mathbf{sl}_2(\mathbf{R})$, the real split form) in $\mathbf{g}$. Precisely, take a nonzero $x$ in $\mathbf{g}_\alpha$. A simple calculation shows that the commutator $h_\alpha = [x, \theta(x)]$ belongs to $\mathbf{a}$. (Additivity of roots for commutators, as in §2.8, shows that $h \in \mathbf{g}_0$, and it is easy to check that $\theta(h_\alpha) = -h_\alpha$.) If $\alpha(h_\alpha) \neq 0$, then one can scale $x$ to get $\alpha(h_\alpha) = 2$, whence $x$, $\theta(x)$, and $h_\alpha$ form a standard basis for $\mathbf{sl}_2$. Further, the possibility that $\alpha(h_\alpha) = 0$ can be shown to contradict the fact that $B(x, \theta(x)) < 0$. (Note the analogy with the Cartan criterion argument, cf. §2.8.)

Thus one gets a copy of $\mathbf{sl}_2(\mathbf{R})$ inside $\mathbf{g}$ for any root $\alpha$ of $\mathbf{a}$. The corresponding copy of $\text{SL}_2(\mathbf{R})$, obtained by exponentiation, will contain elements which normalize $\mathbf{a}$: these elements will induce a reflection in the hyperplane orthogonal to $\alpha$. One concludes that the restricted roots of $\mathbf{a}$ in $\mathbf{g}$ form a root system in the formal sense **[Bour, Crtr, Hump, Serr1]** with Weyl group generated by reflections in the hyperplanes orthogonal to the roots (the root hyperplanes). These hyperplanes divide $\mathbf{a}$ into a collection of convex cones, the Weyl chambers. These are the closures of connected components of $\mathbf{a} \cap \mathbf{p}_{\text{reg}}$. The Weyl chambers are permuted simply transitively by $W$.

A.2.3. *Decompositions of* $G$. Associated to the decompositions (A.2.1.10) and (A.2.2.2) are decompositions of $G$, the connected group with Lie algebra $\mathbf{g}$. Set

$$K = G \cap \text{O}_n = \{k \in G : \theta(k) = k\}.$$

Clearly, the Lie algebra of $K$ is the $\mathbf{k}$ of equation (A.2.1.10). Let $\mathbf{p}$ be the other summand in (A.2.1.10). Let exp be the exponential map from $\mathbf{gl}_n(\mathbf{R})$ to $\text{GL}_n(\mathbf{R})$.

(A.2.3.1) (Cartan decomposition, I). The mapping

$$K \times \mathbf{p} \to G, \qquad (k, x) \to k \exp x$$

is a diffeomorphism. Thus each element $g$ in $G$ has a unique factorization $g = k \exp x$ with $k \in K$, $x \in \mathbf{p}$.

We use the shorthand $G = K \exp \mathbf{p}$ to indicate the state of affairs described in (A.2.3.1).

REMARK. The Cartan decomposition shows that, as topological space, $G \sim K \times \mathbf{R}^m$, where $m = \dim \mathbf{p}$. Hence all interesting topology of $G$, including Betti numbers, homotopy groups, etc., is determined by $K$. Thus $K$ is connected since we have assumed $G$ is connected; also, the argument above did not require $G$ to be closed, so $G$ is closed in $\mathrm{GL}_n(\mathbf{R})$ if and only if $K$ is.

To prove (A.2.3.1), one takes $g$ in $G$ and considers $g^t g$. This is a selfadjoint positive-definite matrix, so it has a unique selfadjoint logarithm $y \in \mathbf{s} \subseteq \mathbf{gl}_n(\mathbf{R})$: $g^t g = \exp y$. Suppose that $y$ is in $\mathbf{p}$. Then $\exp(y/2) = (g^t g)^{1/2}$, and one checks that $k = g(g^t g)^{-1/2}$ is in $K$, so the desired factorization is $g = (g \exp(-y(2))) \exp(y/2)$.

The above argument shows that the desired factorization certainly exists in $\mathrm{GL}_n(\mathbf{R})$ (where it is also known as the polar decomposition or principal value factorization [**Stra**, **Lang2**, **Gaal**]. To show that if $g$ is in $G$, and $g = k \exp x$ is its Cartan decomposition in $\mathrm{GL}_n(\mathbf{R})$, then $x \in \mathbf{p}$, argue as follows. The map $g \to x$ is analytic. For $g$ near the identity, the Inverse Function Theorem implies that $x \in \mathbf{p}$. Since $G$ is connected, we always have $x \in \mathbf{p}$.

REMARK. The Cartan decomposition can be extended to nonconnected subgroups satisfying appropriate conditions (such as being algebraic) [**Knap2**, **Wall2**].

If we combine decomposition (A.2.3.1) with Lemma A.2.2.5, we get another useful decomposition of $G$. Let $\mathbf{a} \subseteq \mathbf{p}$ be a split Cartan subalgebra. Let $A = \exp \mathbf{a}$ be the connected abelian group with Lie algebra $A$. Choose a Weyl chamber (cf. §A.2.2) in $\mathbf{a}$. Denote it by $\mathbf{a}^+$, and set $A^+ = \exp \mathbf{a}^+$.

A.2.3.2 (Cartan decomposition, II). The mapping

$$K \times A^+ \times K \to G, \qquad (k_1, a, k_2) \to k_1 a k_2$$

is surjective.

We indicate this result by writing $G = KA^+K$, or just $G = KAK$.

This decomposition results from combining (A.2.3.1) with Lemma A.2.2.5. If $g = k \exp x$, and $x = \mathrm{Ad}\,\tilde{k}(y)$ for $y \in \mathbf{a}^+$, then $g = k\tilde{k} \exp(a) \tilde{k}^{-1}$, which is decomposition (A.2.3.2) with $k_1 = k\tilde{k}$ and $k_2 = \tilde{k}^{-1}$.

Next consider decompositions of $G$ associated to the root space decomposition (A.2.2.2). Let $\mathbf{a}^+$ be a Weyl chamber in $\mathbf{a}$. By definition, $\mathbf{a}^+$ is a set where each root of $\mathbf{a}$ in $\mathbf{g}$ takes values of only one sign (i.e., either all nonnegative or all nonpositive). Let $\Delta^+$ denote the set of roots which are positive on $\mathbf{a}^+$. Set

(A.2.3.3) $$\mathbf{n}^+ = \sum_{\alpha \in \Delta^+} \mathbf{g}_\alpha.$$

Combining equations (A.2.2.2), (A.2.2.4), and definition (A.2.3.3) gives us the equation

(A.2.3.4) $$\mathbf{g} = \mathbf{k} \oplus \mathbf{a} \oplus \mathbf{n}^+.$$

This is the infinitesimal version of the *Iwasawa decomposition*.

(A.2.3.5) (Iwasawa decomposition). Let $N^+ \subseteq G$ be the connected subgroup whose Lie algebra is $\mathbf{n}^+$. Then $N^+$ is closed in $G$ and

$$\exp : \mathbf{n}^+ \to N^+$$

is a diffeomorphism. Furthermore, the map

$$K \times A \times N^+ \to G, \qquad (k, a, n) \to kan$$

is a surjective diffeomorphism.

This result is usually indicated by the shorthand $G = KAN^+$.

If $G = \mathrm{GL}_n(\mathbf{R})$, we may take $\mathbf{a}$ to be the diagonal matrices and $\mathbf{n}^+$ to be the strictly upper triangular matrices, in which case the Iwasawa decomposition amounts to the Gram-Schmidt orthonormalization procedure [**Hill**, **Stra**]. Also the exponential map on the strictly upper triangular matrices is a polynomial map (of degree $n - 1$) with polynomial inverse.

For general reductive $G \subseteq \mathrm{GL}_n(\mathbf{R})$, if we position $G$ correctly, by conjugation if necessary, we can arrange that the Iwasawa decomposition for $G$ is the same as for $\mathrm{GL}_n(\mathbf{R})$. Indeed we can choose an orthonormal eigenbasis $\{b_j\}_{j=1}^n$ for $\mathbf{R}^n$ consisting of eigenvectors for $\mathbf{a}$, and we can order this eigenbasis by picking an element $x$ in $\mathbf{a}^+$, and requiring that the $x$-eigenvalue of $b_j$ decreases as $j$ increases. Then the commutation relations (A.2.2.2)(ii) show, by a calculation like that of formula (3.5.1.3), that, with respect to the basis $\{b_j\}$, the Lie algebra $\mathbf{n}^+$ consists of strictly upper triangular matrices. Also, $\mathbf{a}$ consists of diagonal matrices, and $\mathbf{k}$ still consists of skew-symmetric matrices.

Now consider $g \in G$, and let $g = kan$ be its Iwasawa decomposition *as an element of* $\mathrm{GL}_n(\mathbf{R})$. The infinitesimal decomposition (A.2.3.4), combined with the Inverse Function Theorem [**Lang2**], implies that for $g$ near enough to the identity, we have $k \in K$, $a \in A$, and $n \in N^+$. Since $G$ is connected and $k$, $a$, and $n$ depend analytically on $g$, it follows that the Iwasawa decomposition for $\mathrm{GL}_n(\mathbf{R})$ provides the Iwasawa decomposition for $G$ also.

We also record, but do not prove, the general version of the Bruhat decomposition (cf. §1.2 for the case of $\mathrm{GL}_n$). Let $M$ be the centralizer in $K$ of $A$. Set $Q_0 = MAN^+$. Let $W$ be the Weyl group of $A$.

(A.2.3.6) (Bruhat decomposition). The group $G$ is a disjoint union of $(N^+, Q_0)$ double cosets with representatives from $W$:

$$G = \bigcup_{w \in W} N^+ w Q_0.$$

This is sometimes abbreviated $G = N^+ W Q_0$.

In the Bruhat decomposition, we can use, instead of $N^+$, the "opposite" unipotent group $N^- = \exp \mathbf{n}^-$, where

$$\mathbf{n}^- = \sum_{\alpha \in \Delta^+} \mathbf{g}_{-\alpha}. \tag{A.2.3.7}$$

It is clear from equation (A.2.2.2) and definitions (A.2.3.3) and (A.2.3.7) that

(A.2.3.8) $$\mathbf{g} = \mathbf{n}^- \oplus \mathbf{m} \oplus \mathbf{a} \oplus \mathbf{n}^+ = \mathbf{n}^- \oplus \mathbf{q}_0,$$

where $\mathbf{q}_0 = \mathbf{m} \oplus \mathbf{a} \oplus \mathbf{n}^+$ is the Lie algebra of $Q_0$. The decomposition (A.2.3.8) implies that the double coset $N^- Q_0 = N^- MAN^+$ is open in $G$. This is the analog for $G$ of the $L - U$ decomposition of §1.1, and the Bruhat decomposition $G = N^- W Q_0$ extends this decomposition to an arbitrary element of $G$.

A.2.4. *Parabolic subgroups.* The group $Q_0 = MAN^+$ figuring in the Bruhat decomposition (A.2.3.6), or any subgroup of $G$ conjugate to $Q_0$, is called a *minimal parabolic subgroup* of $G$. A *parabolic subgroup* of $G$ is any subgroup which contains a minimal parabolic subgroup.

We can construct parabolic subgroups containing $Q_0$ as follows. Let $\mathbf{a}$ be the split Cartan subalgebra used above, and let $\mathbf{a}^+ \subseteq \mathbf{a}$ be the Weyl chamber defining $\mathbf{n}^+$. Thus $\mathbf{a}^+$ is a convex cone, defined as the intersection of certain halfspaces cut out by the root hyperplanes in $\mathbf{a}$. Let $\mathscr{F} = \{\alpha_j\}_{j=1}^r$ be the set of roots which are nonnegative on $\mathbf{a}^+$, and whose kernel intersects $\mathbf{a}^+$ in a set of codimension 1 in $\mathbf{a}$—equivalently $\mathbf{a}^+ \cap \ker \alpha_j$ spans $\ker \alpha_j$. The $\alpha_j$ are called the *fundamental positive roots.* Let $\mathscr{F}_1 \subseteq \mathscr{F}$ be a subset of the fundamental positive roots. Set

(A.2.4.1) $$\mathbf{a}_1 = \bigcap_{\alpha \in \mathscr{F}_1} \ker \alpha = \{x \in \mathbf{a} : \alpha(x) = 0, \text{ all } \alpha \in \mathscr{F}_1\},$$
$$\mathbf{q}_1 = \mathbf{m} + \mathbf{a} + \mathbf{n}^+ + \sum_{\ker \beta \supseteq \mathbf{a}_1} \mathbf{g}_\beta$$

Let $Q_1$ be the subgroup of $G$ generated by $Q_0$ and the exponentials of elements of $Q_1$. Then $Q_1$ is a parabolic subgroup of $G$.

Set

(A.2.4.2) $$\mathbf{n}_1^+ = \sum_{\substack{\beta \in \Delta^+ \\ \mathbf{a}_1 \not\subseteq \ker \beta}} \mathbf{g}_\beta, \qquad N_1^+ = \exp \mathbf{n}_1^+.$$

Then we have the decompositions

(A.2.4.3) $$\mathbf{q}_1 = (\mathbf{q}_1 \cap \theta(\mathbf{q}_1)) \oplus \mathbf{n}_1^+, \qquad Q_1 = (Q_1 \cap \theta(Q_1)) \cdot N_1^+,$$
$$\mathbf{q}_1 \cap \theta(\mathbf{q}_1) = \mathbf{m} \oplus \mathbf{a} \oplus \sum_{\mathbf{a}_1 \subseteq \ker \beta} \mathbf{g}_\beta.$$

We see that $\mathbf{q}_1 \cap \theta(\mathbf{q}_1)$ is the centralizer of $\mathbf{a}_1$ in $\mathbf{g}$, and we can likewise characterize $Q_1 \cap \theta(Q_1)$ as the centralizer of $A_1 = \exp \mathbf{a}_1$ in $G$.

Further decompose $\mathbf{q}_1 \cap \theta(\mathbf{q}_1)$ as follows. Let $\mathbf{m}_1$ denote the Lie algebra generated by $\mathbf{m}$ and the root spaces $\mathbf{g}_\beta$ contained in $\mathbf{q}_1 \cap \theta(\mathbf{q}_1)$. Then it can be checked that $\mathbf{m}_1 \cap \mathbf{a}$ is the span of the elements $[x_\alpha, \theta(x_\alpha)]$, $x_\alpha \in \mathbf{g}_\alpha$, $\alpha \in \mathscr{F}_1$. In particular, $\mathbf{a} = (\mathbf{m}_1 \cap \mathbf{a}) \oplus \mathbf{a}_1$. Thus we have

(A.2.4.4) $$\mathbf{q}_1 \cap \theta(\mathbf{q}_1) = \mathbf{m}_1 \oplus \mathbf{a}_1 \quad \text{and} \quad \mathbf{q}_1 = \mathbf{m}_1 \oplus \mathbf{a}_1 \oplus \mathbf{n}_1^+.$$

This is called the *Langlands decomposition* of $\mathbf{q}_1$ **[GaVa, Knap2, Wall2]**. It is easy to see that $\mathbf{m}_1$ is a reductive Lie subalgebra of $\mathbf{g}$. We can define a corresponding subgroup $M_1$ of $G$ as follows. The group $Q_1 \cap \theta(Q_1)$ has a Cartan decomposition (cf. (A.2.3.1))

$$Q_1 \cap \theta(Q_1) = (Q_1 \cap \theta(Q_1) \cap K)\exp(\mathbf{q}_1 \cap \mathbf{p}).$$

Set

$$M_1 = (Q_1 \cap \theta(Q_1) \cap K)\exp(\mathbf{m}_1 \cap \mathbf{p}). \tag{A.2.4.5}$$

Then the Cartan decomposition plus decompositions (A.2.4.4) tell us that

$$P_1 = M_1 A_1 N_1^+ \tag{A.2.4.6}$$

in the strong sense that the map from $M_1 \times A_1 \times N_1^+$ defined by multiplication to $P_1$ is a diffeomorphisn. The factorization (A.2.4.6) is called the *Langlands decomposition* of $P_1$.

The procedure sketched above constructs $2^r$, where $r = \#(\mathcal{F})$, parabolic subgroups of $G$ containing $P_0$. These are all possible parabolics containing $P_0$. To show this requires a more detailed study of root systems than we wish to give here. Instead we will finish as we started, by looking at $\mathrm{GL}_n$. We will sketch how to see that possibilities for subgroups of $\mathrm{GL}_n$ containing the Borel subgroup of upper triangular matrices are the groups of block upper triangular matrices defined by various partial flags (cf. §1.4). Consider the basis $\{E_{jk}\}_{j,k=1}^n$ of standard matrix units for $\mathbf{gl}_n$. These satisfy the commutation relations

$$[E_{jk}, E_{lm}] = \delta_{kl}E_{jm} - \delta_{jm}E_{lk}.$$

The upper triangular matrices $\mathbf{b}^+$ are the span of the $E_{jk}$ with $j \leq k$. Suppose we add to this another element $x = \sum c_{lm}E_{lm}$. Since the $E_{lm}$'s are eigenvectors for the $\operatorname{ad} E_{jj}$, with distinct eigenvalues, we find that if $c_{lm} \neq 0$, then $E_{lm}$ is in the algebra generated by $\mathbf{b}^+$ and $x$. So take $x = E_{lm}$ for some $l > m$. Taking commutators with $E_{jl}$, $j \leq l$, shows us $E_{jm}$ belongs to the algebra generated by $E_{lm}$ and $\mathbf{b}^+$. Similarly, we must have $E_{lk}$, $k \geq m$, in this algebra. Repeating this process, we find that all $E_{jk}$, $j \leq l$, $k \geq m$, are in the algebra. These span the whole block to the upper right of $E_{lm}$. Next suppose we have two elements $E_{lm}$, $E_{rs}$ which generate overlapping blocks, in the sense that $m < s \leq l < r$. Then from the argument above, we can find $E_{rl}$ in the algebra generated by $\mathbf{b}^+$ and $E_{rs}$. Hence $[E_{rl}, E_{lm}] = E_{rm}$ is in our algebra, and therefore so is the smallest diagonal block containing *both* $E_{lm}$ and $E_{rs}$. Thus we get the general parabolic containing $\mathbf{b}^+$ by adding *disjoint* diagonal blocks. We remark that the calculations sketched above are similar to those used in the context of general root systems.

**Acknowledgments.** In writing this paper I have benefitted from the insights and remarks of many people. I thank James Arthur, Richard Askey, Joseph