

## SZEGÖ LIMIT THEOREMS FOR TOEPLITZ OPERATORS ON COMPACT HOMOGENEOUS SPACES

BY

I. I. HIRSCHMAN, JR., D. S. LIANG AND E. N. WILSON<sup>1</sup>

**ABSTRACT.** Let  $f$  be a real valued integrable function on a compact homogeneous space  $M = K \backslash G$  and  $M_f$  the operator of pointwise multiplication by  $f$ . The authors consider families of Toeplitz operators  $T_{f,P} = PM_fP$  as  $P$  ranges over a net of orthogonal projections from  $L^2(M)$  to finite dimensional  $G$ -invariant subspaces. Necessary and sufficient conditions are given on the net in order that the distribution of eigenvalues of these Toeplitz operators is asymptotic to the distribution of values of  $f$  in the sense of Szegő's classical theorem for the circle. Explicit sequences satisfying these conditions are constructed for all compact Lie groups and for all Riemannian symmetric compact spaces.

**1. Introduction.** Let  $T^1$  be the one dimensional torus, i.e. the group of real numbers modulo  $2\pi$ . For  $n \geq 0$ , let

$$D_n(x) = \sum_{k=0}^n e^{ikx}$$

and

$$(P_n\varphi)(x) = \frac{1}{2\pi} \int_0^{2\pi} D_n(x-y)\varphi(y) dy.$$

The measure  $\mu = \frac{1}{2\pi} dx$  is the normalized Haar measure on  $T^1$  and  $P_n$  projects  $L^1(T^1, \mu)$  onto the  $(n+1)$ -dimensional space spanned by  $\{e^{ikx}\}_{0 \leq k \leq n}$ . In  $L^2(T^1, \mu)$ ,  $P_n$  is the orthogonal projection to this space. For  $f \in L^1(T^1, \mu)$ , let  $M_f$  denote the multiplication operator

$$(M_f\varphi)(x) = f(x)\varphi(x).$$

The Toeplitz operator  $T_{f,n}$  is defined as the restriction to  $P_n L^2(T^1, \mu)$  of  $P_n M_f P_n$ . If  $f$  is real valued,  $T_{f,n}$  is selfadjoint and hence has a complete set  $\{\lambda_{n,k}\}_{0 \leq k \leq n}$  of real eigenvalues. In 1915, Szegő [10] proved that for any  $a \in \mathbf{R}$  for which  $\mu\{x: f(x) = a\} = 0$ ,

$$\mu\{x: f(x) > a\} = \lim_{n \rightarrow \infty} \frac{N^+(a; T_{f,n})}{n+1}$$

---

Received by the editors January 11, 1980 and, in revised form, November 14, 1980.

1980 *Mathematics Subject Classification.* Primary 47B35, 43A85; Secondary 22C05.

*Key words and phrases.* Toeplitz operators, Szegő theorem, compact groups, homogeneous spaces, Lie groups, compact symmetric spaces.

<sup>1</sup>Research supported in part by NSF Grant No. MCS 76-06291.

©1982 American Mathematical Society  
 0002-9947/81/0000-1004/\$07.50

where  $N^+(a; T_{f,n})$  is the number of eigenvalues of  $T_{f,n}$  which are greater than  $a$ . This has the clear interpretation that the distribution of values of  $f$  is recaptured by the asymptotic distribution of eigenvalues of the family of Toeplitz operators  $T_{f,n}$ . It is curious that it was not until 1958 that Grenander and Szegö [3] established the corresponding result for the real line. Subsequently, H. A. Krieger [7] generalized these results to the class of locally compact abelian groups whose dual is compactly generated but not compact.

In this paper we will show that if  $G$  is a compact, not necessarily abelian group, then there exist "Szegö theorems" on coset spaces of  $G$  (in particular, on  $G$  itself) provided certain conditions involving the irreducible representations of  $G$  can be satisfied. In §2, we review standard harmonic analysis tools on compact spaces and end by summarizing the relevant properties of the projection operators which replace the family  $\{P_n\}_{n \geq 0}$  above. §3 begins with a discussion of the type of "Szegö theorem" we have in mind and culminates in Theorem 3.10 with the result announced above. With the aid of a polynomial growth estimate established in §4, §5 shows that the appropriate representation theory conditions are satisfied for many examples, including all compact Lie groups and, more generally, all Riemannian symmetric compact manifolds.

The first two authors initiated consideration of this problem and proved Theorem 3.10 for compact groups and for compact coset spaces  $K \setminus G$  when  $L^1(K \setminus G)$  is an abelian subalgebra of the convolution algebra  $L^1(G, dx)$  (see subsection 2.1 for notations). The third author generalized their method of proof to apply to arbitrary coset spaces of a compact group and supplied the treatment of examples in §5.

## 2. Harmonic analysis on compact homogeneous spaces.

2.1. This section begins by reviewing the harmonic analysis tools which stem from the classical Peter-Weyl Theorem for compact groups. Frequently these tools are referred to collectively as the "Peter-Weyl machinery".

Throughout this section,  $G$  will be a compact Hausdorff group and  $K$  a closed subgroup of  $G$  with  $dx$  (respectively,  $dk$ ) the unique Haar measure of total mass 1 on  $G$  (respectively,  $K$ ). A function  $f$  on  $G$  is said to be  $K$  left-invariant (respectively,  $K$  right-invariant,  $K$  bi-invariant) if  $f(kx) = f(x)$  (respectively,  $f(xk) = f(x)$ ,  $f(kx) = f(xk) = f(x)$ ) for all  $(k, x) \in K \times G$ . Functions  $\tilde{f}$  on the space  $K \setminus G$  of left  $K$  cosets are in one-to-one correspondence with  $K$  left-invariant functions  $f$  on  $G$  via  $\tilde{f}(Kx) = f(x)$ . Although our primary concern is with analysis on coset spaces, it is notationally awkward to maintain the distinction between  $\tilde{f}$  on  $K \setminus G$  and its "pull back"  $f$  on  $G$ . For this reason, we phrase everything in terms of  $K$  left-invariant functions on  $G$  and define  $C(K \setminus G)$  as the subspace of  $K$  left-invariant elements in  $C(G)$ , the space of continuous functions on  $G$ , and  $L^p(K \setminus G)$  ( $p \geq 1$ ) as the subspace of  $K$  left-invariant elements in  $L^p(G, dx)$ .  $L^1(K \setminus G)$  is then a closed subspace of the Banach algebra  $L^1(G, dx)$  where the convolution product of  $f_1$  and  $f_2$  in  $L^1(G, dx)$  is defined by

$$(1) \quad (f_1 * f_2)(x) = \int_G f_1(xy^{-1})f_2(y) dy.$$

The Peter-Weyl machinery amounts to the description of a canonical family of minimal two-sided convolution ideals which yield a Hilbert space direct sum decomposition for  $L^2(G, dx) \subset L^1(G, dx)$ . These ideals are finite dimensional and contained in  $C(G)$ ; when they are thought of as replacements for the one dimensional spaces on the circle  $T = \mathbf{R}/2\pi\mathbf{Z}$  spanned by  $e_k(x) = e^{ikx}$  ( $k \in \mathbf{Z}$ ), a very satisfactory analog of ordinary Fourier analysis is obtained.

2.2. *Definitions and notations.*  $\hat{G}$  denotes the set of equivalence classes of irreducible (hence finite dimensional) unitary representations of  $G$ . For each  $\alpha$ , select a representation  $U_\alpha$  acting on  $(H_\alpha, \langle \cdot, \cdot \rangle)$  which belongs to the class  $\alpha$ . Thus  $H_\alpha$  is a finite dimensional complex vector space,  $\langle \cdot, \cdot \rangle$  is a complex inner product on  $H_\alpha$ , and  $x \rightarrow U_\alpha(x)$  is a continuous homomorphism from  $G$  into the group of operators on  $H_\alpha$  which are unitary with respect to  $\alpha$ . Here as throughout this paper,  $\langle \cdot, \cdot \rangle$  is a generic symbol for inner product with particular inner products determined by the context of a discussion. The *degree of  $\alpha$*  is the dimension of  $H_\alpha$  and is denoted by  $d_\alpha$ .  $L(H_\alpha)$  denotes the space of linear operators on  $H_\alpha$ . Two distinguished elements of  $L(H_\alpha)$  are the identity operator  $I_\alpha$  and, with the usual interpretation of operator-valued integrals, the operator

$$(2) \quad Q_\alpha = \int_K U_\alpha(k) dk.$$

By a trivial computation using translation invariance of  $dk$  and its normalization,  $Q_\alpha$  is the orthogonal projection from  $H_\alpha$  to  $H_{\alpha,K} = \{v \in H_\alpha: U_\alpha(k)v = v \text{ for all } k \in K\}$ . In particular

$$(3) \quad U_\alpha(k)Q_\alpha = Q_\alpha = Q_\alpha U_\alpha(k) \quad \text{for } k \in K.$$

Denote  $\dim H_{\alpha,K} = \text{trace } Q_\alpha$  by  $m_\alpha$ .  $\hat{G}_K$  is defined as the subset of  $\hat{G}$  consisting of those elements  $\alpha$  for which  $m_\alpha \neq 0$ .

For  $\alpha \in \hat{G}$  and  $A \in L(H_\alpha)$ , let

$$(4) \quad f_A(x) = d_\alpha \text{trace}(U_\alpha(x)A)$$

and define

$$(5) \quad C_\alpha = \{f_A: A \in L(H_\alpha)\},$$

$$(6) \quad C_{\alpha,K} = \{f_A: A \in L(H_\alpha) \text{ and } A = AQ_\alpha\}.$$

By (2) and (3),  $C_{\alpha,K}$  is precisely the subspace of  $K$  left-invariant elements in  $C_\alpha$ . Similarly, the subspace of  $K$  bi-invariant elements in  $C_\alpha$  consists of those functions  $f_A$  for which  $AQ_\alpha = A = Q_\alpha A$ . Of paramount importance in what follows are the properties of the  $K$  bi-invariant function

$$(7) \quad \varphi_\alpha(x) = f_{Q_\alpha}(x) = d_\alpha \text{trace}(U_\alpha(x)Q_\alpha).$$

When  $K = \{e\}$ ,  $\varphi_\alpha = d_\alpha \chi_\alpha(x)$  where  $\chi_\alpha(x) = \text{trace } U_\alpha(x)$  is the usual *trace character* of  $U_\alpha$ .

Note that for  $B = \{e_i^{(\alpha)}: 1 \leq i \leq d_\alpha\}$  any orthonormal basis of  $H_\alpha$ , the  $i, j$  entry of the matrix of  $U_\alpha(x)$  relative to  $B$  is the function

$$(8) \quad f_{i,j,\alpha}(x) = \langle U_\alpha(x)e_j^{(\alpha)}, e_i^{(\alpha)} \rangle, \quad x \in G, 1 \leq i, j \leq d_\alpha.$$

By a trivial computation,  $C_\alpha$  is the span in  $C(G)$  of these matrix entries of  $U_\alpha$  and hence may be called the  $\alpha$ -matrix coefficient space of  $G$ . If  $B$  is chosen so that its first  $m_\alpha$  elements constitute a basis of  $H_{\alpha,K}$ , then  $C_{\alpha,K}$  is the span in  $C(K \setminus G)$  of the entries in the first  $m_\alpha$  rows while the  $K$  bi-invariant elements in  $C_\alpha$  are spanned by the entries in the first  $m_\alpha$  rows and columns.

The operator-valued Fourier transform of  $f \in L^1(G, dx)$  is the function  $\hat{f}$  on  $\hat{G}$  which assigns to  $\alpha \in \hat{G}$  the element  $\hat{f}(\alpha) \in L(H_\alpha)$  defined by

$$(9) \quad \hat{f}(\alpha) = \int_G f(x)U_\alpha(x^{-1}) dx.$$

The  $\alpha$ -Fourier component of  $f$  is the function  $f_\alpha \in C_\alpha$  defined using (4) by

$$(10) \quad f_\alpha = f_{\hat{f}(\alpha)}.$$

If  $f \in L^1(K \setminus G)$ , it is easily verified from (2) and (9) that  $\hat{f}(\alpha) = \hat{f}(\alpha)Q_\alpha$  and thus by (6),  $f_\alpha \in C_{\alpha,K}$ ; in particular,  $\hat{f}(\alpha)$  and  $f_\alpha$  are zero if  $\alpha \notin \hat{G}_K$ . For any  $f \in L^1(G, dx)$ , the formal Fourier series expansion of  $f$  is given by

$$(11) \quad f(x) \sim \sum_{\alpha \in \hat{G}} f_\alpha(x) = \sum_{\alpha \in \hat{G}} d_\alpha \text{trace}(U_\alpha(x)\hat{f}(\alpha)).$$

2.3. LEMMA. (i) For  $f^*(x) = \overline{f(x^{-1})}$ ,  $f \rightarrow f^*$  is a conjugate linear anti-automorphism (i.e. product reversing map) of  $L^1(G, dx)$  leaving  $C_\alpha$  invariant for all  $\alpha \in \hat{G}$ .

(ii) For  $f_1, f_2 \in L^2(G, dx)$ ,  $\langle f_1, f_2 \rangle = (f_1 * f_2^*)(e) = (f_2^* * f_1)(e)$ .

(iii) For  $f_1, f_2, f_3 \in L^2(G, dx)$ ,  $\langle f_1 * f_2, f_3 \rangle = \langle f_2, f_1^* * f_3 \rangle = \langle f_1, f_3 * f_2^* \rangle$ .

(iv) For  $f \in L^1(G, dx)$  and  $\alpha \in \hat{G}$ ,  $(f^*)^\wedge(\alpha) = (\hat{f}(\alpha))^*$ , the Hermitian adjoint of  $\hat{f}(\alpha)$ .

(v) For  $f_1, f_2 \in L^1(G, dx)$  and  $\alpha \in \hat{G}$ ,  $(f_1 * f_2)^\wedge(\alpha) = \hat{f}_2(\alpha)\hat{f}_1(\alpha)$ .

PROOF. Routine elementary calculations using only the definitions. Note that here and throughout this paper, we use generically the symbol  $e$  to denote the identity element in a group.

2.4. THEOREM (PETER-WEYL). With the above definitions and notations, the following properties hold:

(i) In the sense of Hilbert-space orthogonal direct sum decompositions,

$$L^2(G, dx) = \bigoplus_{\alpha \in \hat{G}} C_\alpha, \quad L^2(K \setminus G) = \bigoplus_{\alpha \in \hat{G}_K} C_{\alpha,K}.$$

(ii) For  $A \in L(H_\alpha), B \in L(H_\beta), B^*$  the Hermitian adjoint of  $B$ ,

$$(12) \quad \langle f_A, f_B \rangle = \begin{cases} 0 & \text{if } \alpha \neq \beta, \\ d_\alpha \text{trace}(AB^*) & \text{if } \alpha = \beta \end{cases}$$

where the left side of (12) is the inner product in  $L^2(G, dx)$  of the functions defined by (4).

(iii) Using the notations of (4), (9) and (10), if  $f = f_\alpha \in C_\alpha$ , then  $\hat{f}(\beta)$  and  $f_\beta$  are zero for  $\beta \neq \alpha$  while  $\hat{f}(\alpha) = A, f_\alpha = f$ . For arbitrary  $f \in L^2(G, dx)$ , the right side of (11) converges to  $f$  in the sense of  $L^2$ -convergence. If  $f$  is the convolution of two functions in

$L^2(G, dx)$  [in particular, if only finitely many Fourier components  $f_\alpha$  are nonzero], the right side of (11) converges pointwise to  $f$ .

(iv) For  $\alpha \in \hat{G}$ ,  $C_\alpha$  is a minimal two-sided ideal in  $L^1(G, dx)$  with  $\dim C_\alpha = d_\alpha^2$  and  $C_{\alpha,K}$  is a right ideal in  $L^1(G, dx)$  (i.e.  $f * g \in C_{\alpha,K}$  whenever  $f \in C_{\alpha,K}$ ,  $g \in L^1(G, dx)$ ) with  $\dim C_{\alpha,K} = d_\alpha m_\alpha$ .

(v) For  $\alpha \in \hat{G}$ , the orthogonal projection from  $L^2(G, dx)$  to  $C_\alpha$  is given by  $f \rightarrow (d_\alpha \chi_\alpha) * f = f * (d_\alpha \chi_\alpha)$ . The orthogonal projection from  $L^2(G, dx)$  to  $C_{\alpha,K}$  is given by  $f \rightarrow \varphi_\alpha * f$ ; if  $f$  is  $K$  bi-invariant then  $f * \varphi_\alpha = \varphi_\alpha * f$ .

PROOF. The classical Peter-Weyl theorem is the assertion that the functions  $\sqrt{d_\alpha} f_{i,j,\alpha}$  ( $\alpha \in \hat{G}$ ,  $1 \leq i, j \leq d_\alpha$ ) defined in (8) constitute an orthonormal basis for  $L^2(G, dx)$ . See [5, p. 24] for a proof of the classical theorem. In view of the above remarks, (i) is simply a basis-free formulation of the classical theorem. The orthogonality relations (12) follow immediately by expansion of the traces in terms of the matrix entries (8); the dimension statements in (iv) follow from the description above of  $C_\alpha$  and  $C_{\alpha,K}$  in terms of matrix entries. To obtain the first statement in (iii), it is enough to use (12) along with an easy manipulation showing that for  $f = f_A (A \in L(H_\alpha))$  and  $B \in L(H_\alpha)$ ,  $\text{trace}(\hat{f}(\alpha)B^*) = \langle f_A, f_B \rangle / d_\alpha$ . This then implies that for  $f$  arbitrary in  $L^2(G, dx)$ ,  $f_\alpha$  is the component of  $f$  in  $C_\alpha$  relative to the orthogonal decomposition (i) and convergence of (ii) in  $L^2(G, dx)$  is immediate. See [5, p. 25] for the last statement in (iii). The statements about ideals in (iv) follow from (iii), Lemma 2.3(v), and the familiar elementary linear algebra result that  $L(H_\alpha)$  contains no proper two-sided ideals. For  $f \in L^2(G, dx)$ , one checks that  $(d_\alpha \chi_\alpha) * f$  is just  $f_{\hat{f}(\alpha)}$  which, in view of (10) and (iii), is the component of  $f$  in  $C_\alpha$  relative to the decomposition in (ii); a similar computation shows that  $\varphi_\alpha * f = f_{\hat{f}(\alpha)Q_\alpha} \in C_{\alpha,K}$ . These observations easily imply (v).

2.5. *Contragredient and tensor product representations.* For  $\alpha \in \hat{G}$  and  $U_\alpha$  as in 2.2, the *contragredient* of  $U_\alpha$  is the representation  $\overline{U}_\alpha$  defined on the dual space  $H_\alpha^*$  of  $H_\alpha$  by  $\overline{U}_\alpha(x) = {}'U_\alpha(x^{-1})$ . Here, for  $A \in L(H_\alpha)$ ,  $'A \in L(H_\alpha^*)$  is the usual transpose of  $A$ . Trivially for  $J: H_\alpha \rightarrow H_\alpha^*$  the conjugate linear isomorphism obtained from the inner product on  $H_\alpha$  (i.e.  $(Ju)(v) = \langle v, u \rangle$  for  $u, v \in H_\alpha$ ),  $JU_\alpha(x)J^{-1} = \overline{U}_\alpha(x)$ . This means that  $\overline{U}_\alpha$  is an irreducible representation of  $G$  whose matrix entries are complex conjugates of matrix entries of  $U_\alpha$ ; denoting the equivalence class of  $\overline{U}_\alpha$  by  $\bar{\alpha} \in \hat{G}$ ,  $d_{\bar{\alpha}} = d_\alpha$  and  $C_{\bar{\alpha}} = \overline{C_\alpha}$ . In particular,  $\alpha = \bar{\alpha}$  if and only if  $C_\alpha$  is closed under complex conjugation or, in view of Theorem 2.4(v), if and only if  $\chi_\alpha$  is real-valued. The representation  $\overline{U}_\alpha$  is unitary relative to the inner product defined by  $\langle Ju, Jv \rangle = \langle v, u \rangle$  for  $u, v \in H_\alpha$ . It follows that  $Q_{\bar{\alpha}} = {}'Q_\alpha$  (see (2)) and  $\varphi_{\bar{\alpha}} = \overline{\varphi_\alpha}$ , the complex conjugate of  $\varphi_\alpha$  (see (7)).

For  $\alpha, \beta \in \hat{G}$  and  $A \in L(H_\alpha)$ ,  $B \in L(H_\beta)$ ,  $A \otimes B$  denotes the usual tensor product operator acting on  $H_\alpha \otimes H_\beta$ . The tensor product of  $U_\alpha$  and  $U_\beta$  is the representation  $U_\alpha \otimes U_\beta$  defined on  $H_\alpha \otimes H_\beta$  by

$$(U_\alpha \otimes U_\beta)(x) = U_\alpha(x) \otimes U_\beta(x).$$

$U_\alpha \otimes U_\beta$  is unitary relative to the unique inner product on  $H_\alpha \otimes H_\beta$  for which  $\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle = \langle u_1, u_2 \rangle \langle v_1, v_2 \rangle$  for all  $u_1, u_2 \in H_\alpha$ ,  $v_1, v_2 \in H_\beta$ . In general,

$U_\alpha \otimes U_\beta$  is not irreducible. Up to equivalence,  $U_\alpha \otimes U_\beta$  depends only on  $\alpha$  and  $\beta$ ; it is customary to denote its equivalence class by  $\alpha \otimes \beta$ . A class  $\gamma \in \hat{G}$  is said to occur in  $\alpha \otimes \beta$  if there is an irreducible subrepresentation of  $U_\alpha \otimes U_\beta$  which belongs to  $\gamma$ , i.e. is equivalent to  $U_\gamma$ . Denote by  $\Gamma_{\alpha,\beta}$  the subset of  $\hat{G}$  consisting of the elements which occur in  $\alpha \otimes \beta$ . Obviously,  $\Gamma_{\alpha,\beta} = \Gamma_{\beta,\alpha}$ .

2.6. LEMMA. For  $\alpha, \beta \in \hat{G}$ ,  $\Gamma_{\alpha,\beta}$  is finite and for  $\gamma \in \hat{G}$ , the following conditions are equivalent:

- (i)  $\gamma \in \Gamma_{\alpha,\beta}$ ;
- (ii)  $\bar{\beta} \in \Gamma_{\alpha,\bar{\gamma}}$ ;
- (iii)  $\langle \chi_\alpha \chi_\beta, \chi_\gamma \rangle \neq 0$ ;
- (iv) there exists  $f \in C_\alpha, g \in C_\beta$  such that  $(fg)\hat{(\gamma)} \neq 0$ .

PROOF. If  $T$  is a unitary representation of  $G$  on a finite dimensional Hilbert space  $H$ , then  $H$  has an orthogonal direct sum decomposition  $H = \bigoplus_{i=1}^N H_i$  where for  $1 \leq i \leq N$ , the subspace  $H_i$  is invariant and irreducible under  $T$ . Let  $\gamma_i$  be the equivalence class of  $T|_{H_i}$ . For  $\gamma \in \hat{G}$ , let  $n_\gamma$  be the number of indices  $i$  for which  $\gamma_i = \gamma$ . For  $\chi_T(x) = \text{trace } T(x)$  the trace character of  $T$ , it follows that  $\chi_T = \sum_{\gamma \in \hat{G}} n_\gamma \chi_\gamma$ . In view of Theorem 2.4(i), (ii),  $n_\gamma = \langle \chi_T, \chi_\gamma \rangle$ . This shows that the multiplicities  $n_\gamma$  are well defined even though the decomposition  $H = \bigoplus_{i=1}^N H_i$  is not canonical. For  $T = U_\alpha \otimes U_\beta$ ,  $\chi_T = \chi_\alpha \chi_\beta$  and equivalence of (i) and (iii) results. Since  $\langle \chi_\alpha \chi_\beta, \chi_\gamma \rangle = \langle \chi_\alpha \bar{\chi}_\gamma, \bar{\chi}_\beta \rangle = \langle \chi_\alpha \chi_{\bar{\gamma}}, \chi_{\bar{\beta}} \rangle$ , (ii) is equivalent to (i) and thus to (iii). Now let  $f = f_A \in C_\alpha, g = f_B \in C_\beta$  with notations as in (4). For each (i), let  $P_i: H_\alpha \otimes H_\beta \rightarrow H_{\gamma_i}$  be the composition of the orthogonal projection operator  $Q_i$  from  $H_\alpha \otimes H_\beta$  to  $H_i$  with an isometry from  $H_i$  to  $H_{\gamma_i}$  which intertwines  $U_\alpha \otimes U_\beta|_{H_i}$  and  $U_{\gamma_i}$ . Thus

$$(13) \quad P_i^* P_i = Q_i, \quad P_i(U_\alpha \otimes U_\beta)(x) = U_{\gamma_i}(x) P_i \quad \text{for } 1 \leq i \leq N, x \in G.$$

It follows from (13) that

$$(14) \quad f_A(x) f_B(x) = d_\alpha d_\beta \sum_{i=1}^N \text{trace } U_{\gamma_i}(x) P_i(A \otimes B) P_i^*.$$

By Theorem 2.4(iii) and (14), we obtain

$$(15) \quad (f_A f_B)\hat{(\gamma)} = \frac{d_\alpha d_\beta}{d_\gamma} \sum_{\gamma_i = \gamma} P_i(A \otimes B) P_i^*.$$

In particular,  $(f_A f_B)\hat{(\gamma)} = 0$  if  $n_\gamma = 0$ , i.e.  $\gamma \notin \Gamma_{\alpha,\beta}$ . Equivalence of (iv) with (i) follows from (15) and the fact that the vector space of all operators on  $H_\alpha \otimes H_\beta$  is spanned by  $\{A \otimes B: A \in L(H_\alpha), B \in L(H_\beta)\}$ .

2.7. DEFINITION. A function  $\psi \in L^1(G, dx)$  is said to be positive semidefinite if  $\langle \psi * f, f \rangle \geq 0$  for all  $f \in L^2(G, dx)$ . An operator  $A$  on a complex Hilbert space  $(H, \langle \cdot, \cdot \rangle)$  is positive semidefinite if  $\langle Au, u \rangle \geq 0$  for all  $u$ .

2.8. LEMMA. (i) If  $\psi \in L^1(G, dx)$  is positive semidefinite, then  $\psi = \psi^*$  (see Lemma 2.3 for the definition of  $\psi^*$ ).

(ii) For  $\psi \in L^1(G, dx)$ , the following conditions are equivalent:

- (a)  $\psi$  is positive semidefinite;

- (b)  $\psi_\alpha$  (defined as in (10)) is positive semidefinite for all  $\alpha$  in  $\hat{G}$ ;
- (c)  $\hat{\psi}(\alpha) \in L(H_\alpha)$  is positive semidefinite for all  $\alpha$  in  $\hat{G}$ .
- (iii) If  $\psi_1$  and  $\psi_2$  are positive semidefinite in  $L^2(G, dx) \subset L^1(G, dx)$ , then  $\psi = \psi_1\psi_2$  is positive semidefinite.

PROOF. Lemma 2.3(i) implies (i). For  $\psi \in L^1(G, dx)$ ,  $f \in L^2(G, dx)$ , Lemma 2.3(iv) along with Theorem 2.4 and (12) yield

$$\langle \psi * f, f \rangle = \sum_{\alpha \in \hat{G}} \langle \psi_\alpha * f_\alpha, f_\alpha \rangle = \sum_{\alpha \in \hat{G}} d_\alpha \text{trace}(\hat{f}(\alpha)\hat{\psi}(\alpha)(\hat{f}(\alpha))^*).$$

Equivalence of (a)–(c) in (ii) follows from this and the routine linear algebra observation that an operator  $A \in L(H)$  is positive semidefinite if and only if  $\text{trace } BAB^* \geq 0$  for all  $B \in L(H)$ . For (iii), first suppose that  $\psi_1 = f_A \in C_\alpha$ ,  $\psi_2 = f_B \in C_\beta$ . Then  $A = \hat{\psi}_1(\alpha)$ ,  $B = \hat{\psi}_2(\beta)$  are positive semidefinite and in the notation of Lemma 2.6,  $P_i(A \otimes B)P_i^*$  is positive semidefinite for  $i = 1, 2, \dots, N$ . By (15),  $(\psi_1\psi_2)(\gamma)$  is positive semidefinite for all  $\gamma \in \hat{G}$  and hence by (ii),  $\psi_1\psi_2$  is positive semidefinite. By the obvious linearity argument, (iii) follows immediately from this special case if  $\psi_1$  and  $\psi_2$  each have a finite Fourier expansion (12). For  $\psi_1, \psi_2$  arbitrary in  $L^2(G, dx)$ , (iii) is obtained by a routine limit argument involving finite truncations of the Fourier series of  $\psi_1, \psi_2$ .

2.9. For the reader's convenience, the properties of the functions  $\varphi_\alpha$  (see (7)),  $\alpha \in \hat{G}_K$ , will now be summarized.

(i) For  $f \in L^2(K \setminus G)$ , the Fourier component  $f_\alpha$  of  $f$  is equal to  $\varphi_\alpha * f$ ; if  $f$  is  $K$  bi-invariant then  $f_\alpha$  is also equal to  $f * \varphi_\alpha$ . In particular,

$$\varphi_\alpha * \varphi_\beta = \begin{cases} 0, & \alpha \neq \beta, \\ \varphi_\alpha, & \alpha = \beta \end{cases}$$

(see Theorem 2.4(v)).

(ii)  $\varphi_\alpha = \varphi_\alpha^*$  is a positive semidefinite  $K$  bi-invariant function and

$$m_\alpha d_\alpha = \varphi_\alpha(e) = (\varphi_\alpha * \varphi_\alpha^*)(e) = \langle \varphi_\alpha, \varphi_\alpha \rangle$$

(see 2.2 for notations and use Lemmas 2.3 and 2.7).

(iii)  $\varphi_{\bar{\alpha}} = \overline{\varphi_\alpha}$  (see the discussion in 2.5).

(iv)  $\varphi_\alpha \varphi_\beta = \sum_{\gamma \in \Gamma_{\alpha, \beta}} (\varphi_\alpha \varphi_\beta) * \varphi_\gamma = \sum_{\gamma \in \Gamma_{\alpha, \beta}} \varphi_\gamma * (\varphi_\alpha \varphi_\beta)$

(see the discussion in 2.5 and use Lemma 2.6 along with 2.9(i)).

(v) For  $f \in L^2(K \setminus G)$  and  $\beta \in \hat{G}_K$

$$(16) \quad m_\beta d_\beta \|f\|^2 = \sum_{\alpha \in \hat{G}_K} \langle (\varphi_\alpha \overline{\varphi_\beta}) * f, f \rangle.$$

To see why (16) holds, note that by Lemma 2.3 and 2.9(i), (ii),  $\langle (\varphi_\alpha \overline{\varphi_\beta}) * f, f \rangle = \langle \varphi_\beta(f * f^*), \varphi_\alpha \rangle$  is the value at  $e$  of the  $\alpha$ -Fourier component of  $\varphi_\beta(f * f^*)$ . By Theorem 2.4(ii), the Fourier series of  $\varphi_\beta(f * f^*)$  converges pointwise. Since  $m_\beta d_\beta \|f\|^2 = \varphi_\beta(e)(f * f^*)(e)$  by 2.9(ii) and Lemma 2.3, (16) follows.

### 3. The Szegő theorem for compact homogeneous spaces.

3.1. Let  $(M, G)$  be a compact Hausdorff homogeneous space, i.e.  $M$  is a compact Hausdorff space and  $G$  is a compact Hausdorff group which acts transitively on  $M$

in a continuous fashion. Then there is a unique  $G$ -invariant measure  $\mu$  on  $M$  for which  $\mu(M) = 1$ . By  $\text{Rl } L^1(M, \mu)$  we mean the *real-valued* functions in  $L^1(M, \mu)$ . For  $f \in \text{Rl } L^1(M, \mu)$ ,  $M_f$  denotes the operator acting on functions on  $M$  by  $(M_f\varphi)(p) = f(p)\varphi(p)$ . For  $P$  an orthogonal projection on  $L^2(M, \mu)$  with the range of  $P$  a finite dimensional  $G$ -invariant subspace of  $L^2(M, \mu)$ , we define the *Toeplitz operator*  $T_{f,P}$  as the restriction to the range of  $P$  of the operator  $U_{f,P} = PM_fP$ . As will be seen below,  $T_{f,P}$  is well defined. Because  $f$  is real valued,  $T_{f,P}$  is a selfadjoint operator on the finite dimensional space  $PL^2(M, \mu)$  and hence has a real eigenvalue decomposition. Now suppose  $\Sigma(r)$  is a net defined on a directed set  $(\mathfrak{R}, >)$  with values in the collection of finite dimensional  $G$ -invariant subspaces of  $L^2(M, \mu)$ . Let  $P_r$  be the orthogonal projection to  $\Sigma(r)$  and write  $T_{f,r}$  for  $T_{f,P_r}$ . For  $\tau \in \mathbf{R}$ , denote by  $N^+(\tau; T_{f,r})$  the number (with multiplicity) of eigenvalues of  $T_{f,r}$  which are greater than  $\tau$ . We say the *Szegő theorem is valid relative to the net*  $\Sigma$  if for every  $f \in \text{Rl } L^1(M, \mu)$  and every  $\tau \in \mathbf{R}$  for which  $\mu(f^{-1}(\tau)) = 0$ ,

$$(1) \quad \mu(f^{-1}(\tau, \infty)) = \lim_{r \rightarrow \infty} \frac{N^+(\tau; T_{f,r})}{\dim \Sigma(r)}.$$

To make the analysis more tractable, we follow the standard device of identifying  $M$  with a coset space of  $G$  via the choice of a base point  $p_0$  in  $M$ . For minor technical reasons, we prefer to denote the action of  $G$  on the right, i.e. for  $p \in M$  and  $x \in G$ , the action of  $x$  on  $p$  is written  $p \cdot x$ . Then for  $K = \{k \in G: p_0k = p_0\}$ ,  $K$  is a closed subgroup of  $G$  and  $Kx \rightarrow p_0x$  is a homeomorphism from  $K \setminus G$  onto  $M$ . Functions  $\tilde{\varphi}$  on  $M$  are then in one-to-one correspondence with  $K$  left-invariant functions  $\varphi$  on  $G$  via  $\tilde{\varphi}(p_0x) = \varphi(x)$ ; moreover, for  $dx$  the normalized Haar measure on  $G$ , the measure  $\mu$  on  $M$  is given by

$$\int_M \tilde{\varphi}(p)\mu(dp) = \int_G \varphi(x) dx.$$

Using the notations of 2.1,  $L^p(M, \mu)$  then becomes identified with  $L^p(K \setminus G) \subset L^p(G, dx)$  for  $p \geq 1$ . In view of Theorem 2.4(iv), it is natural to look only at nets  $\Sigma$  of the form  $\Sigma(r) = \sum_{\alpha \in \sigma(r)} C_{\alpha,K}$  where  $\sigma(r)$  is a net defined on  $(\mathfrak{R}, >)$  with values in the finite subsets of  $\hat{G}_K$ . Let  $\sigma(r)^\#$  be the dimension of  $\Sigma(r)$ . By Theorem 2.4(iv)-(v),

$$(2) \quad \sigma(r)^\# = \sum_{\alpha \in \sigma(r)} m_\alpha d_\alpha$$

and the projection  $P_r$  to  $\Sigma(r)$  is given by  $P_r f = D_r * f$  where

$$(3) \quad D_r = \sum_{\alpha \in \sigma(r)} \varphi_\alpha.$$

Since  $P_r$  is well defined on  $L^1(K \setminus G)$  as a projection with range in  $C(K \setminus G) \subset L^2(K \setminus G)$ ,  $P_r M_f P_r$  is well defined on  $L^2(K \setminus G) \subset L^1(K \setminus G)$  for  $f \in L^1(K \setminus G)$ .

Denote the Haar measure of any measurable subset  $E$  of  $G$  by  $|E|$ . With these notations and identifications, (1) takes the form

$$(4) \quad |f^{-1}(\tau, \infty)| = \lim_{r \rightarrow \infty} \frac{N^+(\tau; T_{f,r})}{\sigma(r)^\#}$$

for all  $f \in \text{Rl } L^1(K \setminus G)$  provided that  $|f^{-1}(\tau)| = 0$ . We shall begin below by establishing necessary and sufficient conditions on the net  $\sigma$  in order that (4) holds for characteristic functions, then show that under these conditions, (4) survives the familiar journey to arbitrary functions in  $\text{Rl } L^1(K \setminus G)$  via simple functions and nonnegative functions.

When the Hilbert space  $L^2(K \setminus G)$  is separable, arbitrary nets may be dispensed with and replaced by sequences. This will be the case for all of our examples in §5. However, no simplification in our proof results by restricting to the separable case. We do not assume that our nets  $\sigma$  are order preserving in the sense  $\sigma(r_1) \supset \sigma(r_2)$  whenever  $r_1 > r_2$  simply because order considerations are irrelevant in our method of proof. The order preserving property is a natural one, of course, and will be satisfied in all of our examples. With it, our family of Toeplitz operators belongs to the category of abstract Toeplitz operators formulated in [3]. In particular, one has the pleasant property that the largest (respectively, smallest) eigenvalue of  $T_{f,r}$  increases (respectively, decreases) as  $r$  increases; if (4) holds, the limits of these extremal eigenvalues are obviously the essential supremum and essential infimum of  $f$ . Also, the essential infimum of  $f$  is positive, when all of the Toeplitz operators  $T_{f,r}$  are positive definite and this can be verified by the usual minor determinant test with calculations simplified at each stage by the fact that one such minor is a lower order Toeplitz operator. In the classical case (see §1), the  $i, j$  matrix entry of  $T_{f,n}$  is the  $(i - j)$ th Fourier coefficient of  $f$  and hence the Szegő theorem is interpreted as giving substantial information about  $f$  by a discrete set of linear algebra calculations involving its Fourier coefficients. Although explicit calculation of matrix entries for our Toeplitz operators  $T_{f,r}$  is limited by the extent to which tensor product decompositions are known, it is clear from Lemma 2.6 that  $T_{f,r}$  depends only on the Fourier components of  $f$  corresponding to elements  $\gamma$  in the finite subset  $\bigcup_{\alpha, \beta \in \sigma(r)} \Gamma_{\alpha, \beta}$  of  $\hat{G}_K$ . To this extent, the classical interpretation survives. The value of establishing the Szegő theorem on  $K \setminus G$  rather than just on  $G$  should be clear by considering the distinction between analyzing functions on the group  $SO(m)$  and on the sphere in  $\mathbf{R}^m$  regarded as a homogeneous space of  $SO(m)$  (see 5.8). To apply our results in the sense just described, one starts with a spherical harmonic series description of a function on the sphere and then calculates matrix entries for Toeplitz operators with the aid of Clebsch-Gordon coefficients among relevant spherical harmonic representations of  $SO(m)$ . While formidable, this task is much less involved than calculations involving Clebsch-Gordon coefficients for arbitrary representations of  $SO(m)$ .

3.2. LEMMA. *Let  $\sigma(r)$  be any finite subset of  $\hat{G}_K$  with  $\sigma(r)^\#$  and  $D_r$ , defined by (2) and (3). Then,*

- (i)  $(\sigma(r)^\#)^{-1} \int_G |D_r(x)|^2 dx = 1$ ;
- (ii)  $D_r = D_r^* = D_r * D_r$ ;
- (iii)  $0 \leq (\sigma(r)^\#)^{-1} \langle |D_r|^2 * f, f \rangle \leq \langle f, f \rangle$  for all  $f \in L^2(K \setminus G)$ .

PROOF. (i) and (ii) are immediate from 2.9(i), (ii). Since  $\varphi_\alpha$  is positive semidefinite for all  $\alpha \in \hat{G}_K$ , Lemma 2.7 implies that  $\varphi_\alpha \overline{\varphi_\beta}$  is positive semidefinite for  $\alpha, \beta \in \hat{G}_K$

and hence  $|D_r|^2 = \sum_{\alpha, \beta \in \sigma(r)} \varphi_\alpha \overline{\varphi_\beta}$  is positive semidefinite. Using this and 2.9(v), for all  $f \in L^2(K \setminus G)$ ,

$$\begin{aligned} 0 &\leq \langle |D_r|^2 * f, f \rangle \leq \sum_{\beta \in \sigma(r)} \sum_{\alpha \in \hat{G}_K} \langle (\varphi_\alpha \overline{\varphi_\beta}) * f, f \rangle \\ &= \sum_{\beta \in \sigma(r)} m_\beta d_\beta \langle f, f \rangle = \sigma(r)^\# \langle f, f \rangle \end{aligned}$$

and (iii) follows.

3.3. LEMMA. *Let  $\sigma(r)$  be a net with values in the finite subsets of  $\hat{G}_K$ . Then the following conditions are equivalent:*

(i)  $\{(\sigma(r)^\#)^{-1} |D_r|^2 : r \in \mathfrak{R}\}$  is an approximate identity in  $L^2(K \setminus G)$ , i.e.  $\lim_{r \rightarrow \infty} \|(\sigma(r)^\#)^{-1} |D_r|^2 * f - f\|_2 = 0$  for all  $f \in L^2(K \setminus G)$ .

(ii)  $\lim_{r \rightarrow \infty} (\sigma(r)^\#)^{-1} \langle |D_r|^2 * f, g \rangle = \langle f, g \rangle$  for all  $f, g \in L^2(K \setminus G)$ .

(iii)  $\lim_{r \rightarrow \infty} (\sigma(r)^\#)^{-1} \langle |D_r|^2 * f, f \rangle = \langle f, f \rangle$  whenever  $f$  is the characteristic function of a  $K$  left-invariant subset of  $G$ .

(iv) For all  $\gamma \in \hat{G}_K$ ,  $\lim_{r \rightarrow \infty} (\sigma(r)^\#)^{-1} (|D_r|^2 * \varphi_\gamma)(e) = \varphi_\gamma(e)$ .

These equivalent conditions are implied by:

(v) For all  $\gamma \in \hat{G}_K$ ,  $\lim_{r \rightarrow \infty} [\sigma_\gamma^0(r)^\# / \sigma(r)^\#] = 1$  where

$$\sigma_\gamma^0(r) = \{ \alpha \in \sigma(r) : \Gamma_{\alpha, \gamma} \subset \sigma(r) \}, \quad \sigma_\gamma^0(r)^\# = \sum_{\alpha \in \sigma_\gamma^0(r)} m_\alpha d_\alpha$$

(see 2.5 for the definition of  $\Gamma_{\alpha, \gamma}$ ).

PROOF. The implications (i)  $\Rightarrow$  (ii) and (ii)  $\Rightarrow$  (iii) are trivial. For  $\gamma \in \hat{G}_K$ , define  $\psi_{r, \gamma} = (\sigma(r)^\#)^{-1} |D_r|^2 * \varphi_\gamma$ . Since  $|D_r|^2$  is  $K$  bi-invariant,  $\psi_{r, \gamma}$  is the  $\gamma$ -Fourier component of the function  $(\sigma(r)^\#)^{-1} |D_r|^2$  by 2.9(i). Since  $\langle h, \varphi_\gamma \rangle = h(e)$  for all  $h \in C_{\gamma, K}$ , (iv) results from (ii) by taking  $f = g = \varphi_\gamma$ . By Lemma 3.2 for  $f \in L^2(K \setminus G)$  and  $h = \varphi_\gamma * f = f_\gamma$ ,

$$\langle \varphi_\gamma * f, f \rangle - \langle \psi_{r, \gamma} * f, f \rangle = \langle h, h \rangle - (\sigma(r)^\#)^{-1} \langle |D_r|^2 * h, h \rangle \geq 0.$$

Hence,  $\varphi_\gamma - \psi_{r, \gamma}$  is positive semidefinite for all  $r$  and  $\gamma$ . In particular,

$$|\varphi_\gamma(x) - \psi_{r, \gamma}(x)| \leq \varphi_\gamma(e) - \psi_{r, \gamma}(e).$$

Therefore, given (iv),  $\psi_{r, \gamma}$  converges uniformly to  $\varphi_\gamma$  for all  $\gamma$ . If  $f$  has a finite Fourier expansion, it follows that  $(\sigma(r)^\#)^{-1} |D_r|^2 * f = \sum_{\gamma \in \hat{G}_K} \psi_{r, \gamma} * f$  converges uniformly to  $\sum_{\gamma \in \hat{G}_K} \varphi_\gamma * f = f$ . The functions with finite Fourier expansion being dense in  $L^2(K \setminus G)$ , (i) follows.

Finally, we will show that (iii) and (v) separately imply (ii). Define  $Q_r f = f - (\sigma(r)^\#)^{-1} |D_r|^2 * f$  and  $\langle f, g \rangle_r = \langle Q_r f, g \rangle$  for  $f, g \in L^2(K \setminus G)$ . By Lemma 3.2(iii),  $Q_r f$  is a positive semidefinite operator with operator norm  $\leq 1$ . Hence for  $\|f\|_r = \langle f, f \rangle_r^{1/2}$ ,  $\|\cdot\|_r$  is a seminorm on  $L^2(K \setminus G)$  with  $\|f\|_r \leq \|f\|$ ,  $|\langle f, g \rangle_r| \leq \|f\|_r \|g\|_r$  for all  $f, g \in L^2(K \setminus G)$ . Given (iii)  $\lim_{r \rightarrow \infty} \|f\|_r = 0$  whenever  $f$  is a characteristic function in  $L^2(K \setminus G)$  and hence,  $\|\cdot\|_r$  being a seminorm,  $\lim_{r \rightarrow \infty} \|f\|_r = 0$  whenever  $f$  is a simple function in  $L^2(K \setminus G)$ . Since the simple functions are dense in  $L^2(K \setminus G)$  relative to  $\|\cdot\|$ , it follows that  $\lim_{r \rightarrow \infty} \|f\|_r = 0$  for all  $f \in L^2(K \setminus G)$  and then

$\lim_{r \rightarrow \infty} \langle f, g \rangle_r = 0$  for all  $f, g \in L^2(K \setminus G)$ ; this is (ii). Now assume (v). By Lemma 2.6, for  $f \in C_{\gamma, K}$ ,  $(\overline{\varphi_\alpha} \varphi_\beta) * f$  is zero if  $\gamma \notin \Gamma_{\overline{\alpha}\beta}$  and  $\gamma \in \Gamma_{\overline{\alpha}\beta} \Leftrightarrow \overline{\beta} \in \Gamma_{\overline{\alpha}\beta} \Leftrightarrow \beta \in \Gamma_{\alpha\gamma}$ . Hence if  $\alpha \in \sigma_\gamma^0(r)$  and  $f \in C_{\gamma, K}$ ,  $(\overline{\varphi_\alpha} \varphi_\beta) * f = 0$  when  $\beta \notin \sigma(r)$ . Using 2.9(v) and Lemma 3.2(iii), it follows that

$$\begin{aligned} \alpha_\gamma^0(r)^\# \|f\|^2 &= \sum_{\alpha \in \sigma_\gamma^0(r)} m_\alpha d_\alpha \|f\|^2 = \sum_{\alpha \in \sigma_\gamma^0(r)} \sum_{\beta \in \sigma(r)} \langle (\overline{\varphi_\alpha} \varphi_\beta) * f, f \rangle \\ &\leq \langle |D_r|^2 * f, f \rangle \leq \sigma(r)^\# \|f\|^2 \end{aligned}$$

for  $f \in C_{\gamma, K}$ . By (v), this implies  $\lim_{r \rightarrow \infty} \|f\|_r = 0$  whenever  $f \in C_{\gamma, K}$ ,  $\gamma \in \hat{G}_K$ . Then (ii) follows from Theorem 2.4(i) and another application of the properties of  $\|\cdot\|_r$  listed above.

**3.4. REMARKS.** We will eventually show that Lemma 3.3(i)–(iv) are necessary and sufficient for validity of the Szegő theorem (4) on  $(K \setminus G)$  and consequently that Lemma 3.3(v) is sufficient. In practice, conditions (i)–(iv) of Lemma 3.3 are difficult to verify directly. As will be seen below, condition (v) of Lemma 3.3 can be verified in a wide variety of examples. In these examples, there is a concrete sense in which for  $\gamma$  fixed in  $\hat{G}_K$ , the subset  $\Gamma_{\alpha, \gamma}$  forms a neighborhood of  $\alpha$ . With this interpretation,  $\sigma_\gamma^0(r)$  is the “interior” of  $\sigma(r)$  with respect to “ $\gamma$ -neighborhoods” and Lemma 3.3(v) carries the intuitive meaning that for  $r$  large, most of  $\sigma(r)$  should consist of interior points. For the classical case with  $G = T^1$ ,  $K = \{e\}$ ,  $\hat{G}$  is parameterized by  $\mathbf{Z}$  with  $\varphi_k(x) = e^{ikx}$  ( $k \in \mathbf{Z}$ ,  $x \in G$ ) and  $\Gamma_{k, l}$  consists of the single point  $k + l$ . If for  $n \geq 0$ , one defines  $\sigma(n)$  as any consecutive string of  $n + 1$  integers, the classical choice being  $\sigma(n) = \{0, 1, 2, \dots, n\}$ , then Lemma 3.3(v) is just the trivial assertion that for all  $n \in \mathbf{Z}$ ,

$$\frac{n + 1 - |k|}{n + 1} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

For compact Lie groups, it is natural to define  $\sigma(n)$  as the representations whose highest weight has length at most  $n$ ; in §5, we will show that Lemma 3.3(v) then holds. For the classical case and other easy examples, it is possible to verify Lemma 3.3(v) for many choices of  $\sigma$ , including some bizarre ones. A host of Szegő theorems result.

In general,  $|D_r|^2 * \varphi_\gamma$ , the  $\gamma$ -Fourier component of  $|D_r|^2$ , need not be a constant multiple  $C_{r, \gamma}$  of  $\varphi_\gamma$ ; when it is, Lemma 3.2(iv) takes the less forbidding form  $\lim_{r \rightarrow \infty} C_{r, \gamma} = 1$  for all  $\gamma \in \hat{G}_K$ . There are two important special cases where this simplification results. When  $K \setminus G$  is symmetric in the sense that the  $K$  bi-invariant functions in  $G$  commute under convolution, it is easy to prove that  $m_\gamma = 1$  for all  $\gamma \in \hat{G}_K$  and deduce that the functions  $\varphi_\gamma$  span the  $K$  bi-invariant functions; since  $|D_r|^2$  is  $K$  bi-invariant, it follows that  $|D_r|^2 = \sum_{\gamma \in \hat{G}_K} C_{r, \gamma} \varphi_\gamma$ . Alternatively when  $K = \{e\}$ ,  $\varphi_\alpha = d_\alpha \chi_\alpha$  is a central function on  $G$  (i.e.  $\varphi_\alpha(xy) = \varphi_\alpha(yx)$  for all  $x, y \in G$ ) and it follows easily from Theorem 2.4 that the characters  $\chi_\alpha$  form an orthonormal basis for the central functions in  $L^2(G, dx)$ . Since  $|D_r|^2$  is central, we obtain  $|D_r|^2 = \sum_{\gamma \in \hat{G}} C_{r, \gamma} \varphi_\gamma$ . From the discussion above,  $0 \leq C_{r, \gamma} \leq 1$  in all such cases.

3.5. LEMMA. Let  $f \in \text{Rl } L^1(K \setminus G)$ ,  $\sigma(r)$  a net with values in the finite subsets of  $\hat{G}_K$ . Using the notations of 3.1

(i)  $\text{trace } T_{f,r} = \sigma(r) \int_G f(x) dx$ ,

(ii)  $\text{trace } T_{f,r}^2 = \langle |D_r|^2 * f, f \rangle$ ,

(iii)  $\text{trace}(M_g U_{f,r} M_g) = \langle |D_r|^2 * f, g^2 \rangle$  for  $g$  any measurable left  $K$  invariant function on  $G$  which is bounded and real valued.

PROOF. Since  $U_{f,r} = P_r M_f P_r$  agrees with  $T_{f,r}$  on the range of  $P_r$  and vanishes on the orthogonal complement,  $\text{trace } T_{f,r}^k = \text{trace } U_{f,r}^k$  for  $k = 1, 2$ . Using  $P_r \varphi = D_r * \varphi$ ,  $U_{f,r}$  is seen to be a selfadjoint integral operator with the kernel

$$K_{f,r}(x, y) = \int_G D_r(xz^{-1})f(z)D_r(zy^{-1}) dz.$$

Clearly  $K_f$  is a jointly continuous function on  $G \times G$ . Since  $U_{f,r}$  has finite rank, it has only finitely many negative eigenvalues. Thus Mercer's theorem [9, p. 245] applies and with the aid of Lemma 3.2 and an obvious change of variable argument,

$$\text{trace } U_{f,r} = \int_G K_f(x, x) dx = \int_G (|D_r|^2 * f)(x) dx = \sigma(r) \int_G f(x) dx.$$

For (ii),  $U_{f,r}^2$  has the iterated kernel

$$K_{f,r}^{(2)}(x, y) = \int_G K_f(x, u)K_f(u, y) dy = \int_G K_f(x, u) \overline{K_f(y, u)} du$$

and Mercer's theorem, interchange of integration, and Lemma 3.2 yield

$$\begin{aligned} \text{trace } U_{f,r}^2 &= \int_G K_f^{(2)}(x, x) dx \\ &= \int_G \int_G f(z)f(w) \left[ \int_G \int_G D_r(xz^{-1})D_r(zu^{-1})D_r(uw^{-1})D_r(wx^{-1}) dx du \right] dz dw \\ &= \langle |D_r|^2 * f, f \rangle. \end{aligned}$$

For (iii),  $M_g U_{f,r} M_g$  still has finite rank but has the possibly noncontinuous kernel  $g(x)K_{f,r}(x, y)g(y)$  so Mercer's theorem is not applicable. By approximating  $g$  in  $L^2(K \setminus G)$  by a sequence of continuous functions, it is easy to see that the trace of this operator can still be expressed by integrating the kernel along the diagonal. Thus

$$\text{trace}(M_g U_{f,r} M_g) = \int_G \int_G g(x)^2 |D_r|^2(xz^{-1})f(z) dz dx = \langle |D_r|^2 * f, g^2 \rangle.$$

3.6. THEOREM. The equivalent conditions (i)–(iv) of Lemma 3.3 are necessary and sufficient in order that (4) be valid whenever  $f = \chi_\Omega$  is the characteristic function of a left  $K$ -invariant measurable subset  $\Omega$  of  $G$ .

PROOF. When  $f = \chi_\Omega$ , we simplify notation by writing  $M_\Omega$  for  $M_f$  and  $T_{\Omega,r}$  for  $T_{f,r}$ . Since  $T_{f,r} = P_r M_\Omega P_r$  is the product of three orthogonal projections, all of its eigenvalues are between 0 and 1. Both sides of (4) are then zero for  $\tau < 0$  and both sides are 1 for  $\tau > 0$ . For  $\tau = 0$  or 1, (4) is obvious if either  $\Omega$  or its complement is a

null set and otherwise is excluded from consideration since  $\chi_{\Omega}^{-1}(\tau)$  has nonzero measure. It remains to investigate (4) for  $0 < \tau < 1$ , i.e. determine when it is true that

$$(5) \quad |\Omega| = \lim_{r \rightarrow \infty} \frac{N^+(\tau; T_{\Omega,r})}{\sigma(r)^{\#}}$$

for all  $\Omega$  and  $\tau$  in  $(0, 1)$ .

Fix  $\Omega$  and let  $(\lambda_{j,r})_{1 \leq j \leq \sigma(r)^{\#}}$  be the eigenvalues (with multiplicity) of  $T_{\Omega,r}$ . By Lemma 3.5 and 3.2,

$$\sum_{j=1}^{\sigma(r)^{\#}} \lambda_{j,r} = \sigma(r)^{\#} |\Omega|, \quad \sum_{j=1}^{\sigma(r)^{\#}} \lambda_{j,r}^2 = \sigma(r)^{\#} |\Omega| (1 - \varepsilon(r))$$

where

$$\varepsilon(r) = \frac{1}{|\Omega|} \left\{ \langle \chi_{\Omega}, \chi_{\Omega} \rangle - \frac{1}{\sigma(r)^{\#}} \langle |D_r|^2 * \chi_{\Omega}, \chi_{\Omega} \rangle \right\}.$$

Condition (iii) of Lemma 3.3 states that  $\lim_{r \rightarrow \infty} \varepsilon(r) = 0$ . Assuming this,  $\varepsilon(r) < \frac{1}{2}$  for  $r$  sufficiently large. Define

$$\begin{aligned} N_1(r) &= \{j: \lambda_{j,r} < \varepsilon(r)^{1/2}\}, \\ N_2(r) &= \{j: \varepsilon(r)^{1/2} \leq \lambda_{j,r} \leq 1 - \varepsilon(r)^{1/2}\}, \\ N_3(r) &= \{j: \lambda_{j,r} > 1 - \varepsilon(r)^{1/2}\} \end{aligned}$$

and let  $N_i^{\#}(r)$  be the cardinality of  $N_i(r)$  for  $i = 1, 2, 3$ . Then

$$\begin{aligned} \frac{1}{2} \varepsilon(r)^{1/2} N_2^{\#}(r) &\leq \sum_{j \in N_2(r)} \lambda_{j,r} (1 - \lambda_{j,r}) \\ &\leq \sum_{j=1}^{\sigma(r)^{\#}} \lambda_{j,r} - \lambda_{j,r}^2 = \sigma(r)^{\#} |\Omega| \varepsilon(r) \end{aligned}$$

which yields

$$(6) \quad N_2^{\#}(r) = o(\sigma(r)^{\#}).$$

Then

$$\begin{aligned} \sum_{j \in N_3(r)} \lambda_{j,r} &= \sum_{j=1}^{\sigma(r)^{\#}} \lambda_{j,r} - \sum_{j \in N_1(r) \cup N_2(r)} \lambda_{j,r} \\ &= \sigma(r)^{\#} |\Omega| + o(\sigma(r)^{\#}) \end{aligned}$$

and

$$(1 - \varepsilon(r)^{1/2}) N_3^{\#}(r) \leq \sum_{j \in N_3(r)} \lambda_{j,r} \leq N_3^{\#}(r)$$

so

$$(7) \quad N_3^{\#}(r) = \sigma(r)^{\#} |\Omega| + o(\sigma(r)^{\#}).$$

Given  $0 < \tau < 1$ , for  $r$  sufficiently large,  $\varepsilon(r)^{1/2} < \tau < 1 - \varepsilon(r)^{1/2}$ . Combining (6) and (7), we obtain

$$N^+(\tau; T_{\Omega,r}) = N_3^\#(r) + o(\sigma(r)^\#) = \sigma(r)^\# |\Omega| + o(\sigma(r)^\#)$$

and this is precisely (5).

Conversely, suppose we have (5) for all  $\tau$ . For  $0 < \tau_1 < \tau_2 < 1$ , we have  $N^+(\tau_1; T_{\Omega,r}) - N^+(\tau_2; T_{\Omega,r}) = o(\sigma(r)^\#)$  which implies

$$\sigma(r)^\# |\Omega| \varepsilon(r) = \sum_{j=1}^{\sigma(r)^\#} \lambda_{j,r} - \lambda_{j,r}^2 \leq (\tau_1 + 1 - \tau_2)\sigma(r)^\# + o(\sigma(r)^\#).$$

Taking  $\tau_1$  arbitrarily close to 0 and  $\tau_2$  close to 1, it follows that  $\lim_{r \rightarrow \infty} \varepsilon(r) = 0$ , i.e. condition (iii) of Lemma 3.3 holds for  $f = \chi_\Omega$ .

3.7. DEFINITION. For  $A$  a compact selfadjoint operator on a Hilbert space  $H$  and  $\tau > 0$ ,  $N^+(\tau; A)$  denotes the number with multiplicity of eigenvalues of  $A$  which are greater than  $\tau$ .

We do not know a reference for the following elementary results.

3.8. LEMMA. Let  $A$  and  $B$  be compact selfadjoint operators on  $H$ .

(i)  $N^+(\tau; A) \geq N^+(\tau; B)$  for all  $\tau > 0$  if  $A \geq B$ , i.e. if  $A - B$  is a positive semidefinite operator;

(ii)  $N^+(\tau_1 + \tau_2; A + B) \leq N^+(\tau_1; A) + N^+(\tau_2; B)$  for all  $\tau_1, \tau_2 > 0$ ;

(iii)  $N^+(\tau + \delta; A) - N^+(\delta; -B) \leq N^+(\tau; A + B) \leq N^+(\tau - \delta; A) + N^+(\delta; B)$  whenever  $\tau > \delta > 0$ ;

(iv)  $N^+(\tau; \delta A) = N^+(\tau/\delta; A)$  for  $\tau, \delta > 0$ .

PROOF. Define  $P_{\tau,A}$  as the orthogonal projection on  $H$  to  $H_{\tau,A}$ , the sum of all eigenspaces of  $A$  corresponding to eigenvalues greater than  $\tau$ . Since  $A$  is compact,  $H_{\tau,A}$  is finite dimensional and  $N^+(\tau; A) = \dim H_{\tau,A}$ . Then if  $v$  is nonzero in  $H_{\tau,A}$ ,  $\langle Av, v \rangle > \tau \langle v, v \rangle$ . Conversely, if  $\langle Av, v \rangle > \tau \langle v, v \rangle$ , then  $P_{\tau,A}v \neq 0$ . For  $A \geq B$  and  $v$  nonzero in  $H_{\tau,B}$ ,  $\langle Av, v \rangle \geq \langle Bv, v \rangle > \tau \langle v, v \rangle$  implies  $P_{\tau,A}v \neq 0$ . Thus,  $P_{\tau,A}$  is one-to-one on  $H_{\tau,B}$  and (i) follows. For (ii), if  $v$  is nonzero in  $H_{\tau_1+\tau_2; A+B}$ ,

$$\langle Av, v \rangle + \langle Bv, v \rangle = \langle (A + B)v, v \rangle > \tau_1 \langle v, v \rangle + \tau_2 \langle v, v \rangle$$

which means either  $\langle Av, v \rangle > \tau_1 \langle v, v \rangle$  or  $\langle Bv, v \rangle > \tau_2 \langle v, v \rangle$ . The map  $v \rightarrow (P_{\tau_1,A}v, P_{\tau_2,B}v)$  is then one-to-one from  $H_{\tau_1+\tau_2; A+B}$  into the outer direct sum of  $H_{\tau_1,A}$  and  $H_{\tau_2,B}$  and (ii) follows. Statement (iii) is a consequence of (ii) with  $A = (A + B) + (-B)$  and (iv) follows trivially from Definition 3.7.

3.9. LEMMA. Let  $\Omega_1, \dots, \Omega_m$  be a family of mutually disjoint  $K$  left-invariant subsets of  $G$  and  $\sigma(r)$  a net satisfying Lemma 3.3(i)–(iv). For  $f = \chi_{\Omega_j}$ , write  $M_j$  for  $M_j$  and  $U_{j,r}$  for  $U_{j,r}$ . Then for  $\tau > 0$  and  $1 \leq i, j, k \leq m$ ,

(i)

$$N^+(\tau; \pm M_i U_{j,r} M_i) = o(\sigma(r)^\#) \quad \text{if } i \neq j;$$

(ii)

$$N^+(\tau; \pm (M_i U_{j,r} M_k + M_k U_{j,r} M_i)) = o(\sigma(r)^\#) \quad \text{if } i \neq j \text{ or } j \neq k;$$

(iii)

$$N^+(\tau; M_j U_{j,r} M_j) = \begin{cases} o(\sigma(r)^\#) & \text{if } \tau > 1, \\ \sigma(r)^\# |\Omega_j| + o(\sigma(r)^\#) & \text{if } \tau < 1. \end{cases}$$

PROOF. (i) From Lemma 3.5(iii) and Lemma 3.3(ii)

$$\text{trace}(M_i U_{j,r} M_i) = \langle |D_r|^2 * \chi_{\Omega_i}, \chi_{\Omega_j} \rangle = \langle \chi_{\Omega_i}, \chi_{\Omega_j} \rangle + o(\sigma(r)^\#) = o(\sigma(r)^\#).$$

Since  $M_i U_{j,r} M_i$  is a positive semidefinite operator of finite rank,  $N^+(\tau; -M_i U_{j,r} M_i) = 0$  and  $N^+(\tau; M_i U_{j,r} M_i) \leq \tau^{-1} \text{trace}(M_i U_{j,r} M_i) = o(\sigma(r)^\#)$ .

(ii) For  $\rho > 0$  define

$$A_\pm = \rho^{-1} M_i P_r M_j \pm \rho M_k P_r M_j.$$

Then  $A_\pm A_\pm^* \geq 0$  where the  $\pm$  signs are either both “+” or both “-”. Expanding out and using  $M_j^2 = M_j$ ,  $U_{j,r} = P_r M_j P_r$ ,

$$(8) \quad \rho^{-2} M_i U_{j,r} M_i + \rho^2 M_k U_{j,r} M_k \geq \pm (M_i U_{j,r} M_k + M_k U_{j,r} M_i).$$

If neither  $i$  nor  $k$  is equal to  $j$ , (ii) follows from (8) along with (i) and Lemma 3.8. Suppose then that  $i = j \neq k$  and choose  $\rho$  such that  $\rho^2 \tau / 2 > 1$ . Since  $\|M_j U_{j,r} M_j\| \leq 1$ ,  $N^+(\rho^2 \tau / 2; M_j U_{j,r} M_j) = 0$ . By Lemma 3.8 and (8),

$$\begin{aligned} N^+(\tau; \pm (M_j U_{j,r} M_k + M_k U_{j,r} M_j)) \\ \leq N^+(\tau/2; \rho^{-2} M_j U_{j,r} M_j) + N^+(\tau/2; \rho^2 M_k U_{j,r} M_k) \\ = N^+(\tau/(2\rho^2); M_k U_{j,r} M_k) = o(\sigma(r)^\#). \end{aligned}$$

(iii) Let  $\Omega'_j$  be the complement of  $\Omega_j$  in  $G$  and write  $M'_j$  for  $M_{\chi_{\Omega'_j}}$ . Since  $M_j + M'_j$  is the identity operator on  $L^2(K \setminus G)$ ,

$$U_{j,r} = M_j U_{j,r} M_j + V_{j,r}$$

where

$$V_{j,r} = M'_j U_{j,r} M'_j + M_j U_{j,r} M'_j + M'_j U_{j,r} M_j.$$

By (i) and (ii) along with Lemma 3.8(ii),  $N^+(\tau; \pm V_{j,r}) = o(\sigma(r)^\#)$ . Hence by Lemma 3.8(iii), for  $\tau > \delta > 0$ ,

$$(9) \quad \begin{aligned} N^+(\tau + \delta; U_{j,r}) + o(\sigma(r)^\#) \\ \leq N^+(\tau; M_j U_{j,r} U_j) \leq N^+(\tau - \delta; U_{j,r}) + o(\sigma(r)^\#). \end{aligned}$$

For  $\tau > 1$ , we have already noted in the proof of (ii) that  $N^+(\tau; M_j U_{j,r} M_j) = 0$ . For  $\tau < 1$ , we recall from Theorem 3.6 that  $N^+(\tau; U_{j,r}) = \sigma(r)^\# |\Omega_j| + o(\sigma(r)^\#)$ . Choosing  $\delta$  small enough so that  $\tau + \delta < 1$ , (iii) follows from this and (9).

3.10. THEOREM. *The Szegő theorem on  $K \setminus G$  is valid relative to a net  $\sigma$  if and only if  $\sigma$  satisfies the equivalent conditions (i)–(iv) of Lemma 3.3.*

PROOF. The necessity of the conditions in Lemma 3.3 has already been established in Theorem 3.6. We shall prove that these conditions are sufficient by first using

Theorem 3.6 to show that (4) is valid for simple functions and then using approximation by simple functions to obtain (4) in general.

If  $f \in \text{Rl } L^1(K \setminus G)$ ,  $a \in \mathbf{R}$ , and  $(f + a)(x) = f(x) + a$ ,  $T_{f+a,r} = T_{f,r} + aI_r$  where  $I_r$  is the identity operator on  $P_r L^2(K \setminus G)$ . Then, for any  $\tau \in \mathbf{R}$ ,  $N^+(\tau; T_{f+a,r}) = N^+(\tau - a; T_{f,r})$  and  $|(f + a)^{-1}(\tau, \infty)| = |f^{-1}(\tau - a, \infty)|$ . It follows that (4) holds for  $f$  if and only if it holds for  $f + a$ . In particular, validity of (4) for positive-valued functions implies its validity for functions bounded from below.

Let  $f \in \text{Rl } L^1(K \setminus G)$  be a simple function i.e.  $f = \sum_{j=1}^n a_j \chi_{\Omega_j}$  where  $a_1, \dots, a_n$  are real numbers and  $\Omega_1, \dots, \Omega_n$  are pairwise disjoint measurable  $K$ -left invariant subsets of  $G$  whose union is  $G$ . By the above remarks, we may assume without loss of generality that  $a_j > 0$  for all  $j$ . Then both sides of (4) are zero for  $\tau \leq 0$ . Fix  $\tau > 0$  distinct from  $a_1, a_2, \dots, a_n$ . Then  $N^+(\tau; T_{f,r}) = N^+(\tau; U_{f,r})$ . Using the notational conventions of Lemma 3.9

$$U_{f,r} = \left( \sum_{i=1}^n M_i \right) U_{f,r} \left( \sum_{k=1}^n M_k \right) = \sum_{j=1}^n a_j M_j U_{j,r} M_j + W_r$$

where

$$W_r = \sum_{\substack{i \neq j \text{ or} \\ j \neq k}} a_j M_i U_{j,r} M_k.$$

By Lemmas 3.8(iii) and 3.9(ii), (iii),

$$\begin{aligned} N^+ \left( \tau + \delta; \sum_{j=1}^n M_j U_{j,r} M_j \right) + o(\sigma(r)^\#) &\leq N^+(\tau; U_{f,r}) \\ &\leq N^+ \left( \tau - \delta; \sum_{j=1}^n a_j M_j U_{j,r} M_j \right) + o(\sigma(r)^\#) \end{aligned}$$

whenever  $0 < \delta < \tau$ . Choose  $\delta$  small enough so that none of the  $a_j$ 's are in the interval  $[\tau - \delta, \tau + \delta]$ . By Lemmas 3.8(iv), 3.9(i), and the fact that  $\sum_{j=1}^n a_j M_j U_{j,r} M_j$  is an operator direct sum (i.e. the product of any two distinct summands is zero),

$$\begin{aligned} N^+ \left( \tau \pm \delta; \sum_{j=1}^n a_j M_j U_{j,r} M_j \right) &= \sum_{j=1}^n N^+ \left( \frac{\tau \pm \delta}{a_j}; M_j U_{j,r} M_j \right) \\ &= \sum_{a_j > \tau} \sigma(r)^\# |\Omega_j| + o(\sigma(r)^\#). \end{aligned}$$

Combining the last two equations, we obtain (4).

Next let  $f \in L^1(K \setminus G)$  with  $f(x) > 0$  a.e. Then both sides of (4) are zero for  $\tau \leq 0$ . Fix  $\tau > 0$  with  $|f^{-1}(\tau)| = 0$  and choose a monotonically increasing sequence of simple functions  $g_n$  ( $n \geq 1$ ) with  $f(x) = \lim_{n \rightarrow \infty} g_n(x)$  a.e. and  $|g_n^{-1}(\tau)| = 0$  for all  $n$ . By Lemma 3.8(i),  $N^+(\tau; T_{f,r}) \geq N^+(\tau; T_{g_n,r})$  since  $f \geq g_n$  obviously implies  $T_{f,r} \geq T_{g_n,r}$ . Since (4) holds for  $g_n$ , we obtain

$$\begin{aligned} (10) \quad \lim_{r \rightarrow \infty} \frac{N^+(\tau; T_{f,r})}{\sigma(r)^\#} &\geq \lim_{n \rightarrow \infty} \lim_{r \rightarrow \infty} \frac{N^+(\tau; T_{g_n,r})}{\sigma(r)^\#} \\ &= \lim_{n \rightarrow \infty} |g_n^{-1}(\tau, \infty)| = |f^{-1}(\tau, \infty)|. \end{aligned}$$

Now choose any sequence  $\delta_k$  ( $k \geq 1$ ) with  $\delta_k \rightarrow 0^+$  such that  $0 = |f^{-1}(\tau - \delta_k)| = |g_n^{-1}(\tau - \delta_k)|$  for all  $n$  and  $k$ . Since  $f - g_n \geq 0$ ,  $N^+(\delta_k; T_{f-g_n, r}) \leq 1/\delta_k \text{ trace}(T_{f-g_n, r})$ . It follows by Lemmas 3.5(i) and 3.8(iii) that

$$\overline{\lim}_{r \rightarrow \infty} \frac{N^+(\tau; T_{f, r})}{\sigma(r)^\#} \leq |g_n^{-1}(\tau - \delta_k, \infty)| + \frac{1}{\delta_k} \|f - g_n\|_1$$

for all  $k$  and  $n$ . By first letting  $n \rightarrow \infty$  and then letting  $k \rightarrow \infty$ , we obtain

$$\overline{\lim}_{r \rightarrow \infty} \frac{N^+(\tau; T_{f, r})}{\sigma(r)^\#} \leq \lim_{k \rightarrow \infty} |f^{-1}(\tau - \delta_k, \infty)| = |f^{-1}[\tau, \infty)]| = |f^{-1}(\tau, \infty)|$$

and conclude from (10) that (4) holds.

Finally let  $f$  be arbitrary in  $\text{RI } L^1(K \setminus G)$  and define for  $n = 1, 2, \dots$

$$f_n(x) = \max\{f(x), -n\}.$$

Fix any real number  $\tau$  for which  $|f^{-1}(\tau)| = 0$ . For large  $n$   $f_n^{-1}(\tau, \infty) = f^{-1}(\tau, \infty)$ . From the last paragraph and the remarks at the outset of the proof, (4) holds for  $f_n$ . Since  $f_n \geq f$ , Lemma 3.8(i) yields

$$(11) \quad \overline{\lim}_{r \rightarrow \infty} \frac{N^+(\tau; T_{f, r})}{\sigma(r)^\#} \leq |f^{-1}(\tau, \infty)|.$$

To obtain a reverse estimate, we choose  $\delta_k \rightarrow 0^+$  with  $|f^{-1}(\tau + \delta_k)| = 0$  for  $k = 1, 2, \dots$ . Then, for all  $k, n$ ,

$$N^+(\tau + \delta_k; T_{f_n, r}) \leq N^+(\tau; T_{f, r}) + N^+(\delta_k; T_{f_n - f, r})$$

by the obvious extension of Lemma 3.8(iii) to arbitrary  $\tau$  for operators on a finite dimensional Hilbert space. As above,  $N^+(\delta_k; T_{f_n - f, r}) \leq \delta_k^{-1} \sigma(r)^\# \|f_n - f\|_1$ . Hence,

$$\begin{aligned} \overline{\lim}_{r \rightarrow \infty} \frac{N^+(\tau; T_{f, r})}{\sigma(r)^\#} &\geq \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \{|f_n^{-1}(\tau + \delta_k, \infty)| - \delta_k^{-1} \|f_n - f\|_1\} \\ &= \lim_{k \rightarrow \infty} |f^{-1}(\tau + \delta_k, \infty)| = |f^{-1}(\tau, \infty)| \end{aligned}$$

and (4) follows from this and (11).

**4. A polynomial growth estimate.**

4.1. Let  $V$  be an  $m$ -dimensional real vector space. Recall that a subset  $\mathcal{L}$  of  $V$  is said to be a *lattice in  $V$*  if there exists a basis  $v_1, v_2, \dots, v_m$  of  $V$  such that  $\xi = \sum_{i=1}^m x_i v_i \in \mathcal{L}$  if and only if  $x_1, \dots, x_m$  are integers. Any such basis is then said to be a *lattice basis* for  $\mathcal{L}$ . A subset  $\mathcal{C}$  of  $V$  is said to be an *open convex cone* if  $\mathcal{C}$  is a convex set which is open with respect to the standard topology in  $V$  and has the property that whenever  $\xi \in \mathcal{C}$ , then the ray  $\{t\xi : t > 0\}$  is contained in  $\mathcal{C}$ . The following elementary proposition will be used in §5 to insure that condition 3.3(v) can be verified for many examples.

4.2. PROPOSITION. Let  $V$  be an  $m$ -dimensional real vector space,  $\|\cdot\|$  a norm on  $V$ ,  $\mathcal{L}$  a lattice in  $V$ ,  $\mathcal{C}$  an open convex cone in  $V$ , and  $f$  a real-valued polynomial function on  $V$  with  $f(x) > 0$  for  $x \in \bar{\mathcal{C}}$ , the closure of  $\mathcal{C}$ . For  $n > 0$ , define  $\sigma(n) = \{x \in \bar{\mathcal{C}} \cap \mathcal{L} : \|x\| \leq n\}$ . Then for any  $c > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{\sum_{x \in \sigma(n-c)} f(x)}{\sum_{x \in \sigma(n)} f(x)} = 1.$$

PROOF. By choosing a lattice basis for  $\mathcal{L}$  in  $V$ , all the objects at hand can be transferred to similar objects in  $\mathbf{R}^m$ . Hence it suffices to prove the proposition for the case when  $V = \mathbf{R}^m$  and  $\mathcal{L}$  is the lattice consisting of all  $m$ -tuples with integer components. Let  $dx$  be Lebesgue measure in  $\mathbf{R}^m$ , i.e.  $dx = dx_1, dx_2, \dots, dx_m$  for  $x = (x_1, x_2, \dots, x_m)$ . For  $a = (a_1, a_2, \dots, a_m)$  arbitrary and  $D_a = \{x_1, x_2, \dots, x_m : 0 \leq x_i - a_i < 1\}$ , the Lebesgue measure of  $D_a$  is 1. Consequently, with  $B_n = \{x \in \bar{\mathcal{C}} : \|x\| \leq n\}$ , for any function  $F$  on  $\mathbf{R}^m$ ,  $\sum_{x \in \sigma(n)} F(x)$  is a Riemann sum approximation to  $\int_{B_n} F(x) dx$ . For  $S = \{\xi \in \mathbf{R}^m : \|\xi\| = 1\}$ ,  $dx$  is expressed in polar coordinates  $x = r\xi$  ( $r = \|x\|$ ,  $\xi \in S$ ) by  $r^{m-1} dr d\omega(\xi)$  where  $\omega$  is a measure on  $S$ . Thus, for any  $F$ ,

$$\int_{B_n} F(x) dx = \int_{S \cap \bar{\mathcal{C}}} \left( \int_0^n F(r\xi) r^{m-1} dr \right) d\omega(\xi).$$

Now let  $N$  be the degree of  $f$  and  $h$  the unique homogeneous polynomial of degree  $N$  such that  $g = f - h$  has degree  $\leq N - 1$ . Then there exists a constant  $A$  such that  $|g(x)| \leq A \|x\|^{N-1}$  for all  $x$  and

$$(1) \quad \int_{B_n} g(x) dx \leq A \omega(S \cap \bar{\mathcal{C}}) \int_0^n r^{N+m-2} dr = O(n^{N+m-1}).$$

For  $x \in \mathcal{C}$  and  $t > 0$ ,

$$0 < f(tx) = t^N h(x) + g(tx) \leq t^{N-1} (th(x) + A)$$

and hence  $h(x) \geq 0$ . It follows that  $\int_{S \cap \mathcal{C}} h(\xi) d\omega(\xi) > 0$ . Hence

$$(2) \quad \int_{B_n} h(x) dx = \left( \int_0^n r^{N+m-1} dr \right) \int_{S \cap \mathcal{C}} h(\xi) d\omega(\xi) = Kn^{N+m}$$

with  $K > 0$ . By (1) and (2),  $\int_{B_n} f(x) dx = Kn^{N+m} + o(n^{N+m})$  as  $n \rightarrow \infty$ . By routine but tedious arguments which we shall omit, it can be shown that  $\sum_{x \in \sigma(n)} f(x) = Kn^{N+m} + o(n^{N+m})$  as well. The proposition is then immediate.

### 5. Examples.

5.1. In this section, the major result is Corollary 5.7 asserting that the Szegő theorem holds on all connected compact Riemannian symmetric spaces. By the familiar device of identifying a group  $G$  with  $\Delta \setminus G \times G$  where  $\Delta = \{(g_1, g_2) \in G \times G : g_1 = g_2\}$ , it then follows that the Szegő theorem holds for all connected compact Lie groups. For harmonic analysis purposes, this description is needlessly awkward and we therefore treat the group case separately in Theorem 5.4. This has the extra advantage of not burdening the group case by the very technical description of representations in Theorem 5.6. The objects discussed in subsection 5.2 and the

results of subsection 5.3 are all standard. A thorough treatment assuming no background in Lie theory is given in [1].

5.2. Let  $G$  be a compact connected Lie group with Lie algebra  $\mathfrak{g}$  and  $\mathfrak{t}$  a fixed maximal abelian subalgebra of  $\mathfrak{g}$ . Denote by  $\exp$  the usual exponential map from  $\mathfrak{g}$  onto  $G$ . Then  $T = \exp \mathfrak{t}$  is a maximal torus in  $G$  and  $\mathfrak{L} = \{X \in \mathfrak{t}: \exp X = e\}$  is a lattice in  $\mathfrak{t}$ . Let  $\mathfrak{t}^*$  be the dual space of  $\mathfrak{t}$  and  $\mathfrak{L}^* = \{\lambda \in \mathfrak{t}^*: \lambda(H) \in \mathbf{Z} \text{ for all } H \in \mathfrak{L}\}$  the dual lattice. For  $\lambda \in \mathfrak{L}^*$ ,  $\psi_\lambda(\exp H) = e^{2\pi i \lambda(H)}$  defines a continuous homomorphism from  $T$  to  $T^1$ , the unique circle in  $\mathbf{C}$ , and every continuous homomorphism from  $T$  to  $T^1$  is of the form  $\psi_\lambda$  for some  $\lambda \in \mathfrak{L}^*$ . For  $U$  any finite dimensional unitary representation of  $G$  on  $(H, \langle \cdot, \cdot \rangle)$ ,  $H$  has an orthonormal basis  $e_j$  ( $1 \leq j \leq d = \dim H$ ) consisting of simultaneous eigenvectors for the commuting unitary operators  $U(\exp H)$ ,  $H \in \mathfrak{t}$ . It follows that there are elements  $\lambda_1, \lambda_2, \dots, \lambda_d$  in  $\mathfrak{L}^*$  for which the restriction to  $T$  of the trace character  $\chi_U(x) = \text{trace } U(x)$  is given by  $\chi_U|_T = \sum_{j=1}^d \psi_{\lambda_j}$ ; these elements are called the *weights* of  $U$  relative to  $\mathfrak{t}$ . The *roots* of  $\mathfrak{g}$  relative to  $\mathfrak{t}$  are the nonzero weights of the adjoint representation of  $G$  on  $\mathfrak{g}^{\mathbf{C}}$ . This representation is defined by using complex linearity to extend to  $\mathfrak{g}^{\mathbf{C}} = \mathfrak{g} + i\mathfrak{g}$  the operators  $\text{Ad } x$  ( $x \in G$ ) defined on  $\mathfrak{g}$  by  $\text{Ad } x(X) = (\frac{d}{dt})_{t=0} x(\exp tX)x^{-1}$ . The collection of roots of  $\mathfrak{g}$  relative to  $\mathfrak{t}$  will be denoted by  $\Delta$ .

The *Weyl group*  $W$  associated with the pair  $(G, T)$  is defined by  $W = N(T)/T$  where  $N(T)$  is the normalizer of  $T$  in  $G$ . For  $w = nT \in W$ ,  $\text{Ad } n$  maps  $\mathfrak{L}$  into  $\mathfrak{L}$  and hence  $w$  acts on  $\mathfrak{L}^*$  by  $(w \cdot \lambda)(H) = \lambda(\text{Ad } n^{-1}H)$ . The action of  $W$  on  $\mathfrak{L}^*$  preserves the set of weights of any representation of  $G$ . It is convenient to define a partial ordering  $>$  on  $\mathfrak{L}^*$  by  $\lambda > \mu$  if  $\mu$  is contained in the convex hull of  $W \cdot \lambda$  but  $\lambda$  is not contained in the convex hull of  $W \cdot \mu$ . A *Weyl chamber* is a connected component  $\mathcal{C}$  of  $\mathfrak{t} - \bigcup_{\xi \in \Delta} \ker \xi$ . A root  $\xi$  is said to be *positive relative* to  $\mathcal{C}$  if  $\xi(H) > 0$  for some (hence all)  $H \in \mathcal{C}$ ; the collection of positive roots is denoted by  $\Delta^+(\mathcal{C})$ . For  $\mathcal{C}$  a fixed Weyl chamber,  $\delta = \delta(\mathcal{C})$  is defined by

$$\delta = \frac{1}{2} \sum_{\xi \in \Delta^+(\mathcal{C})} \xi.$$

Now fix an inner product  $\langle \cdot, \cdot \rangle$  on  $\mathfrak{g}$  relative to which  $\text{Ad } x$  is orthogonal for all  $x \in G$ . The restriction of  $\langle \cdot, \cdot \rangle$  to  $\mathfrak{t}$  then defines an isomorphism of  $\mathfrak{t}^*$  and  $\mathfrak{t}$  via  $\lambda \leftrightarrow H_\lambda$  where  $\lambda(H) = \langle H, H_\lambda \rangle$  for all  $H \in \mathfrak{t}$ . An inner product on  $\mathfrak{t}^*$  may then be defined by  $\langle \lambda, \mu \rangle = \langle H_\lambda, H_\mu \rangle$  for  $\lambda, \mu \in \mathfrak{t}^*$ . Using this inner product, the dual Weyl chamber  $\mathcal{C}^*$  is defined by

$$\mathcal{C}^* = \{\lambda \in \mathfrak{t}^*: H_\lambda \in \mathcal{C}\} = \{\lambda \in \mathfrak{t}^*: \langle \lambda, \xi \rangle > 0 \text{ for all } \xi \in \Delta^+(\mathcal{C})\}.$$

For  $\delta$  defined as above, it can be shown that  $2\langle \delta, \xi \rangle / \langle \xi, \xi \rangle$  is a positive integer for each  $\xi \in \Delta^+(\mathcal{C})$ . In particular,  $\delta \in \mathcal{C}^*$ .

For application of this machinery to a description of  $\hat{G}$ , recall from 2.4 that  $\alpha \in \hat{G}$  is completely described by its trace character  $\chi_\alpha$ . It can be shown that every  $x \in G$  is of the form  $x = tyt^{-1}$  for some  $y \in G$ ,  $t \in T$  and hence  $\chi_\alpha(x) = \chi_\alpha(t)$ . Thus  $\alpha$  and  $\chi_\alpha$  are determined by  $\chi_\alpha|_T$  and hence by the weights of  $\alpha$  relative to  $\mathfrak{t}$ . The collection of weights is a union in  $\mathfrak{L}^*$  of Weyl group orbits. It can be shown that for  $\bar{\mathcal{C}}^*$  a fixed dual Weyl chamber with  $\mathcal{C}^*$  its closure in  $\mathfrak{t}^*$ , every Weyl group orbit in  $\mathfrak{L}^*$  contains

precisely one point in  $\bar{\mathcal{C}}^*$ , i.e.  $\bar{\mathcal{C}}^* \cap \mathfrak{L}^*$  parametrizes the  $W$ -orbits in  $\mathfrak{L}^*$ . The crux of the famous Cartan-Weyl parametrization of  $\hat{G}$  by “highest weights” is that for  $\alpha \in \hat{G}$ , one  $W$ -orbit in its collection of weights dominates all others in the sense of the above partial ordering on  $\mathfrak{L}^*$ . The precise statement is given in Theorem 5.3; for convenience below, the well known Weyl dimension formula is also stated.

5.3. THEOREM. *Let  $G$  be a compact connected Lie group with  $T = \exp t$  a maximal torus in  $G$ ,  $\mathcal{C}^* \subset t^*$  a dual Weyl chamber. Using the notations of 5.2:*

(i) *For  $\alpha \in \hat{G}$ , there exists a unique element  $\lambda(\alpha) \in \mathfrak{L}^* \cap \bar{\mathcal{C}}^*$  such that  $\chi_\alpha|_T = \sum_{w \in W} w \cdot \Psi_{\lambda(\alpha)} + \sum_{\mu < \lambda(\alpha)} n_\mu \Psi_\mu$  for some nonnegative integers  $n_\mu$ .  $\lambda(\alpha)$  is called the highest weight of  $\alpha$ .*

(ii) *For  $\lambda \in \mathfrak{L}^* \cap \bar{\mathcal{C}}^*$ , there is a unique class  $\alpha \in \hat{G}$  for which  $\lambda = \lambda(\alpha)$ .*

(iii) *For  $\lambda = \lambda(\alpha) \in \mathfrak{L}^* \cap \bar{\mathcal{C}}^*$ , the degree  $d_\alpha$  of  $\alpha$  is given by*

$$(1) \quad d_\alpha = \prod_{\xi \in \Delta^+(\mathcal{C})} \frac{\langle \lambda + \delta, \xi \rangle}{\langle \delta, \xi \rangle}.$$

PROOF. See [1, 6].

5.4. THEOREM. *For  $G$  any compact Lie group, a net  $\sigma$  in  $\hat{G}$  can be chosen so that the Szegő theorem is true on  $G$ , i.e. (4) of §3 is satisfied with  $K = \{e\}$ .*

PROOF. First assume that  $G$  is connected. Using the notations of Theorem 5.3, define for  $n = 1, 0, \dots$ ,

$$(2) \quad \sigma(n) = \{ \alpha \in \hat{G} : \|\lambda(\alpha)\| \leq n \}$$

where  $\|\cdot\|$  is any norm on  $t^*$  invariant under the action of the Weyl group. Note that  $\sigma(n)$  is finite since the lattice  $\mathfrak{L}^*$  can contain only finitely many points with length less than or equal to  $n$ . Now fix  $\alpha, \gamma \in \hat{G}$  with  $\lambda = \lambda(\alpha)$ ,  $\eta = \lambda(\gamma)$ . It follows from Klimyk’s formula [6, p. 142] that  $\beta \in \Gamma_{\alpha, \gamma}$  (see 2.5) implies  $\mu = \lambda(\beta)$  is of the form  $\mu = \lambda + \nu$  where  $\nu$  is a weight of  $\gamma$ . But then either  $\nu \in W \cdot \eta$  or  $\nu < \eta$  and in both cases  $\|\nu\| \leq \|\eta\|$  so

$$\|\lambda(\beta)\| = \|\mu\| \leq \|\lambda\| + \|\nu\| = \|\lambda(\alpha)\| + \|\lambda(\gamma)\|$$

whenever  $\beta \in \Gamma_{\alpha, \gamma}$ . In particular, if  $\alpha \in \sigma(n - \|\lambda(\gamma)\|)$ , then  $\Gamma_{\alpha, \gamma} \subset \sigma(n)$ . In the notation of Lemma 3.3(v),  $\sigma_\gamma^\circ(n) \supset \sigma(n - \|\lambda(\gamma)\|)$  for all  $\gamma \in \hat{G}$ . Now define  $f$  on  $t^*$  by

$$f(\lambda) = \left[ \prod_{\xi \in \Delta^+(\mathcal{C})} \frac{\langle \lambda + \delta, \xi \rangle}{\langle \delta, \xi \rangle} \right]^2.$$

Clearly  $f$  is a polynomial function on  $t^*$  which is positive on  $\bar{\mathcal{C}}^*$ . In the notations of §2,  $m_\alpha = d_\alpha$  since  $K = \{e\}$  and thus by (1),  $f(\lambda) = m_\alpha d_\alpha = d_\alpha^2$  for  $\lambda = \lambda(\alpha)$ . By Proposition 4.2 we conclude that Lemma 3.3(v) is valid when  $\sigma$  is defined by (2) and hence by Theorem 3.10 the Szegő theorem holds on  $G$ .

Now let  $\tilde{G}$  be a nonconnected compact Lie group with  $G$  the connected component of the identity in  $\tilde{G}$ . Then  $\tilde{G}/G$  is both compact and discrete, hence finite. Fix representatives  $p_1, \dots, p_m$  for the cosets of  $G$  in  $\tilde{G}$ . A function  $f$  on  $\tilde{G}$  is naturally identified with an  $m$ -tuple  $(f_1, f_2, \dots, f_m)$  of functions on  $G$  where  $f_j(x) = f(xp_j)$  for

$x \in G$ . The normalized Haar measures  $d\tilde{x}$  on  $\tilde{G}$  and  $dx$  on  $G$  are related by

$$\int_{\tilde{G}} f(\tilde{x}) d\tilde{x} = \frac{1}{m} \sum_{j=1}^m \int_G f_j(x) dx.$$

Hence  $L^2(\tilde{G}, d\tilde{x})$  is identified in a natural way with the outer direct sum of  $m$  copies of  $L^2(G, dx)$ . Let  $\sigma(n)$  be as in (2) with  $\Sigma(n) = \sum_{\alpha \in \sigma(n)} C_\alpha$  the associated subspace of  $L^2(G, dx)$ . Define  $\tilde{\Sigma}(n) \subset L^2(\tilde{G}, d\tilde{x})$  as the outer direct sum of  $m$  copies of  $\Sigma(n)$ . The orthogonal projection operator  $\tilde{P}(n)$  from  $L^2(\tilde{G}, d\tilde{x})$  to  $\tilde{\Sigma}(n)$  is then the outer direct sum of  $m$  copies of the orthogonal projection operator  $P(n): L^2(G, dx) \rightarrow \Sigma(n)$ . For  $f \in \text{Rl } L^1(\tilde{G}, d\tilde{x})$  with  $f \approx (f_1, \dots, f_m)$ , the Toeplitz operator  $\tilde{T}_{f,n} = \tilde{P}(n)M_f\tilde{P}(n)$  is then the direct sum of the Toeplitz operators  $T_{f_j,n} = P_nM_{f_j}P_n$  ( $1 \leq j \leq m$ ). Since equation (1) of 3.1 holds for the operators  $T_{f_j,n}$  relative to  $\Sigma(n)$ , it follows trivially that this equation also holds for  $\tilde{T}_{f,n}$  relative to  $\tilde{\Sigma}(n)$ . The subspaces  $\tilde{\Sigma}(n)$  are easily seen to be invariant under right  $\tilde{G}$  translations. If one chooses the norm  $\| \cdot \|$  as arising as in 5.2 from the restriction to  $\mathfrak{t}$  of an inner product on  $\mathfrak{g}$  relative to which  $\text{Ad } \tilde{x}: \mathfrak{g} \rightarrow \mathfrak{g}$  is orthogonal for all  $x \in \tilde{G}$ , then for  $\alpha \in \hat{G}$  with highest weight  $\lambda(\alpha)$  and  $1 \leq i \leq m$ , the highest weight of the irreducible representation of  $G$  with trace character  $x \rightarrow \chi_\alpha(p_i x p_i^{-1})$  is easily seen to have the same length as  $\lambda(\alpha)$ . It follows that the spaces  $\tilde{\Sigma}(n)$  are invariant under left  $\tilde{G}$ -translations as well. By Theorem 2.4, there is a finite subset  $\tilde{\sigma}(n)$  of  $(\hat{G})^\wedge$  such that  $\tilde{\Sigma}(n) = \sum_{\tilde{\alpha} \in \tilde{\sigma}(n)} C_{\tilde{\alpha}}$ . Consequently, one has the Szegő theorem on  $\tilde{G}$  in the more natural form (4) of 3.1. For most standard examples, it is easy to describe  $\tilde{\sigma}(n)$  explicitly in terms of  $\sigma(n)$ ; a general description is awkward.

5.5. We now wish to discuss the Szegő theorem in the form (1) of §3 for the pairs  $(M, G)$  where  $M$  is a compact Riemannian manifold which is symmetric in the sense of Cartan and  $G$  is the group of isometries on  $M$ . For simplicity, we shall treat only the case where  $M$  is connected and  $G$  is the connected component of the identity in the group of isometries on  $M$ ; more general cases can then be handled by arguments similar to those in the proof of Theorem 5.4. It is well known (e.g. see Helgason [4]) that the pairs  $(M, G)$  are precisely those which have a coset realization  $(K \backslash G, G)$  where  $G$  is a compact connected Lie group,  $K$  is a closed subgroup of  $G$ ,  $G$  acts effectively on  $K \backslash G$  (i.e.  $K$  contains no nontrivial subgroups which are normal in  $G$ ), and there exists an involutive automorphism  $\theta$  of  $G$  for which

$$(3) \quad (K_\theta)_e \subset K \subset K_\theta$$

where  $K_\theta = \{x \in G: \theta(x) = x\}$  and  $(K_\theta)_e$  is the connected component of the identity in  $K$ . We therefore treat only coset spaces of this form and seek a description of a sequence  $\sigma(n)$  for which Lemma 3.3(v) is valid. The description of  $\hat{G}$  in Theorem 5.3 applies to any maximal torus  $T$  and Weyl chamber  $\mathcal{C}$ . For a satisfactory description of  $\hat{G}_K$ , special choices of  $T$  and  $\mathcal{C}$  are required.

Let  $\mathfrak{g}$  be the Lie algebra of  $G$ . Then  $\mathfrak{g}$  has the familiar Cartan decomposition  $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$  (vector space direct sum) where  $\mathfrak{k}$  (respectively,  $\mathfrak{p}$ ) is the  $+1$  (respectively,

-1) eigenspace of the differential  $\theta_*$  of  $\theta$ . By (3),  $\mathfrak{k}$  is the Lie algebra of  $K$ . Since  $\theta_*$  is an automorphism of  $\mathfrak{g}$ ,  $\mathfrak{k}$  and  $\mathfrak{p}$  satisfy the Lie bracket relations

$$(4) \quad [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}.$$

Choose a maximal abelian subspace  $\mathfrak{t}_1$  in  $\mathfrak{p}$  and then choose a maximal abelian subspace  $\mathfrak{t}_2$  in

$$\mathfrak{m} = \{X \in \mathfrak{k} : [X, H] = 0 \text{ for all } H \in \mathfrak{t}_1\}.$$

Let  $\mathfrak{t} = \mathfrak{t}_1 + \mathfrak{t}_2$ . By (4),  $\mathfrak{t}$  is a maximal abelian subalgebra of  $\mathfrak{g}$  and hence  $T = \exp \mathfrak{t}$  is a maximal torus in  $G$ . Any linear functional  $\xi \in \mathfrak{t}^*$  is then of the form  $\xi = \xi_1 + \xi_2$  for  $\xi_j \in \mathfrak{t}_j^*$  ( $j = 1, 2$ ) in the sense that  $\xi(H_1 + H_2) = \xi_1(H_1) + \xi_2(H_2)$  for  $H_j \in \mathfrak{t}_j$  ( $j = 1, 2$ ). As in 5.2, let  $\Delta \subset \mathfrak{t}^* - \{0\}$  be the roots of  $\mathfrak{g}$  relative to  $\mathfrak{t}$ . For  $j = 1, 2$ , let  $\Delta_j \subset \mathfrak{t}_j^*$  be the collection of elements  $\xi_j$  for which there exists  $\xi \in \Delta$  of the form  $\xi = \xi_1 + \xi_2$ . Then  $\Delta_1$  and  $\Delta_2$  are finite sets. Hence there exists  $H_1 \in \mathfrak{t}_1, H_2 \in \mathfrak{t}_2$  such that  $|\xi_1(H_1)| > |\xi_2(H_2)| > 0$  for all nonzero  $\xi_1 \in \Delta_1, \xi_2 \in \Delta_2$ . Let  $H = H_1 + H_2, \mathcal{C}$  the Weyl chamber of  $\mathfrak{t}$  containing  $H$  and  $\mathcal{C}^* \subset \mathfrak{t}^*$  the corresponding dual Weyl chamber.

Let  $\mathcal{L} \subset \mathfrak{t}$  and  $\mathcal{L}^* \subset \mathfrak{t}^*$  be as defined in 5.2. For  $\mathcal{C}^*$  as above, define

$$\mathcal{L}_K^* = \{\lambda \in \mathcal{L}^* : \lambda(H) \in \mathbf{Z} \text{ whenever } H \in \mathfrak{t} \text{ and } \exp H \in K\}.$$

Note that  $\mathcal{L}_K^*$  is a subgroup of the additive group  $\mathcal{L}^*$  and if  $\lambda \in \mathcal{L}_K^*$ , then  $\lambda$  vanishes on  $\mathfrak{t}_2 \subset \mathfrak{k}$ . However the converse need not be true since there may be points in  $\mathfrak{t}_1 \subset \mathfrak{p}$  whose exponential is in  $K$ .

5.6. THEOREM. *Using the notations of 5.5,  $\mathcal{L}_K^* \cap \bar{\mathcal{C}}^*$  parametrizes  $\hat{G}_K$  in the sense of Theorem 5.3, i.e. for  $\alpha \in \hat{G}$ ,  $\alpha$  is in  $\hat{G}_K$  if and only if its highest weight  $\lambda(\alpha) \in \bar{\mathcal{C}}^*$  is in  $\mathcal{L}_K^*$ .*

PROOF. The authors do not know a reference to this result in the form stated although it is presumably familiar to experts in the subject. For the case when  $G$  is simply connected, the theorem is due to E. Cartan and S. Helgason; a proof appears in the book by Garth Warner [11, pp. 210–213]. We shall give only the modifications needed to extend Warner’s proof to the present generality.

Since  $G$  is compact, its Lie algebra  $\mathfrak{g}$  is the Lie algebra direct sum of a compact semisimple Lie algebra  $\mathfrak{g}_0$  and an abelian algebra  $\mathfrak{s}$ . Let  $G_0$  be a connected, simply connected Lie group with Lie algebra  $\mathfrak{g}_0$ ; by Weyl’s theorem [4, p. 123],  $G_0$  is compact. By standard covering group arguments, there exists a torus  $S$  with Lie algebra  $\mathfrak{s}$  and a continuous homomorphism  $\pi$  from  $G_1 = G_0 \times S$  (Lie group direct product) onto  $G$  whose kernel  $D$  is a finite subgroup of the center of  $G_1$ . Define  $K_1 = \pi^{-1}(K)$ . The sets  $\hat{G}_K$  and  $(\hat{G}_1)_{K_1}$  are then in one-to-one correspondence. To see this note that  $U \leftrightarrow U_1 = U \circ \pi$  is a one-to-one correspondence between representations  $U$  of  $G$  and representations  $U_1$  of  $G_1$  for which  $U_1(d)$  is the identity operator for all  $d \in D$ . This correspondence respects equivalence of representations and hence defines an injection  $\alpha \rightarrow \alpha_1$  from  $\hat{G}$  into  $\hat{G}_1$ . Trivially if  $\alpha \in \hat{G}_K$  then  $\alpha_1 \in (\hat{G}_1)_{K_1}$ . Conversely, if  $\beta$  is arbitrary in  $(\hat{G}_1)_{K_1}$  with  $U_\beta$  an element of the class  $\beta$ , Schur’s Lemma implies that for  $d \in D$ ,  $U_\beta(d)$  is a scalar multiple  $c(d)$  of the

identity operator. Since  $D \subset K_1$  and there are nonzero vectors fixed by  $U_\beta(K_1)$  for all  $k_1 \in K_1$ , the scalars  $c(d)$  are 1 and it follows that  $\beta$  is of the form  $\alpha_1$  for some  $\alpha \in \hat{G}_K$ . Since  $G_1$  is the direct product of  $G_0$  and the abelian group  $S$ ,  $\hat{G}_1$  is the Cartesian product of  $\hat{G}_0$  and the familiar character group  $\hat{S} \approx \mathbf{Z}^n$  ( $n = \dim \mathfrak{s}$ ) in the sense that a representation  $U_1$  of  $G_1$  is irreducible if and only if

$$U_1(g_0, s) = \Psi(s)U_0(g_0)$$

where  $\Psi \in \hat{S}$  and  $U_0$  is an irreducible representation of  $G_0$ . In view of these observations, it remains to describe the set  $\mathcal{C}^* \cap \mathcal{L}_K^* \subset \mathfrak{t}^*$  in terms of an analogous set arising from  $G_0$  and  $S$ .

First note that since  $\theta_*$  is an automorphism of  $\mathfrak{g}$ , the center  $\mathfrak{s}$  of  $\mathfrak{g}$  and the derived subalgebra  $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}_0$  are  $\theta_*$  invariant and hence are the direct sums of their intersections with the  $\theta_*$ -eigenspaces  $\mathfrak{k}$  and  $\mathfrak{p}$ . But  $\exp(\mathfrak{k} \cap \mathfrak{s})$  is a connected subgroup of the center of  $G$  and is contained in  $K$ ; since  $G$  acts effectively on  $K \setminus G$  (see 5.5), this subgroup is trivial and  $\mathfrak{k} \cap \mathfrak{s} = \{0\}$ . Hence  $\mathfrak{s} \subset \mathfrak{p}$  and  $\mathfrak{k} \subset \mathfrak{g}_0$ ; moreover,  $\mathfrak{t}_1$ , being maximal abelian in  $\mathfrak{p}$ , must contain  $\mathfrak{s}$ . Thus  $\mathfrak{t} = \mathfrak{t}_0 + \mathfrak{s}$  where  $\mathfrak{t}_0 = (\mathfrak{t}_1 \cap \mathfrak{g}_0) + \mathfrak{t}_2$  is a maximal abelian subalgebra of  $\mathfrak{g}_0$ . The groups  $G_1$  and  $G$  both have Lie algebra  $\mathfrak{g}$ . Exponential maps are distinguished by writing  $\exp_G$  (respectively,  $\exp_{G_1}$ ) for the exponential map from  $\mathfrak{g}$  onto  $G$  (respectively,  $G_1$ ). Then  $\exp_G = \pi_0 \exp_{G_1}$ . Since  $G_0$  is simply connected, there is a unique involutive automorphism  $\theta_0$  of  $G_0$  whose differential is the restriction of  $\theta_*$  to  $\mathfrak{g}_0$ .  $\theta_0$  may then be extended to an involutive automorphism  $\theta_1$  of  $G_1$  defined by  $\theta_1(g_0, s) = (\theta_0(g_0), s^{-1})$ . Let  $K_0 = \{g_0 \in G_0: \theta_0(g_0) = g_0\}$ . The triple  $(G_0, K_0, \theta_0)$  satisfies (3); by one of the fundamental results on compact symmetric spaces,  $G_0$  simply connected implies  $K_0$  connected [4, p. 272]. Since the construction of  $\theta_1$  implies that  $\pi \circ \theta_1 = \theta \circ \pi$ , it follows from (3) for  $(G, K, \theta)$  that  $K_0 \subset K_1$ . However  $K, K_0$ , and  $K_1$  all have Lie algebra  $\mathfrak{k}$  so  $K_0$  is the connected component of the identity in  $K_1$  and  $K_1/K_0$  is finite. Using the familiar Cartan decomposition  $G_1 = K_0 (\exp_{G_1} \mathfrak{t}_1) K_0$ , elements  $H_1, H_2, \dots, H_r$  in  $\mathfrak{t}_1$  may be selected so that  $K_1$  is the disjoint union of the cosets  $K_0 \exp_{G_1} H_j$  ( $1 \leq j \leq r$ ). The normalized Haar measures  $dk_0$  on  $K_0$  and  $dk_1$  on  $K_1$  are then related by

$$(5) \quad \int_{K_1} f(k_1) dk_1 = \frac{1}{r} \sum_{j=1}^r \int_{K_0} f(k_0 \exp_{G_1} H_j) dk_0.$$

From the above we have  $\mathfrak{t} = \mathfrak{t}_0 + \mathfrak{s}$  and hence in an obvious sense,  $\mathfrak{t}^* = \mathfrak{t}_0^* + \mathfrak{s}^*$  and the roots of  $\mathfrak{g}$  with respect to  $\mathfrak{t}$  all belong to  $\mathfrak{t}_0^*$ . The dual Weyl chamber  $\mathcal{C}^*$  selected in 5.5 is then  $\mathcal{C}_0^* + \mathfrak{s}^*$  where  $\mathcal{C}_0^* \subset \mathfrak{t}_0^*$  is a dual Weyl chamber relative to  $(\mathfrak{g}_0, \mathfrak{t}_0)$  constructed from  $(G_0, K_0, \theta_0)$  in exactly the same fashion that  $\mathcal{C}^*$  was constructed from  $(G, K, \theta)$ . Moreover with  $\mathcal{L}_0 \subset \mathfrak{t}_0$  and  $(\mathcal{L}_0^*)_{K_0}$  defined from  $(G_0, K_0, \theta_0)$  as in 5.2 and 5.5, it is clear that for any  $\lambda \in \mathcal{L}_K^*, \lambda_0 = \lambda|_{\mathfrak{t}_0}$  is in  $(\mathcal{L}_0^*)_{K_0}$ . In Warner's proof, it is shown that if  $\lambda_0 \in \mathcal{C}_0^* \cap (\mathcal{L}_0^*)_{K_0}$  with  $U_{\alpha_0}$  an irreducible representative of  $G_0$  having highest weight  $\lambda_0$  (see Theorem 5.3), then for  $Q_{\alpha_0}$  the orthogonal projection defined by (2) of subsection 2.1 and  $e_{\lambda_0}$  a nonzero weight vector for the weight  $\lambda_0$  (i.e.  $U_{\alpha_0}(t)e_{\lambda_0} = \psi_{\lambda_0}(t)e_{\lambda_0}$  for all  $t \in T_0$ ), then  $Q_{\alpha_0}e_{\lambda_0} \neq 0$ .

Now suppose  $\alpha \in \hat{G}$  has highest weight  $\lambda(\alpha) = \lambda \in \mathfrak{L}_K^* \cap \bar{\mathfrak{C}}^*$ , and  $U_\alpha$  belongs to the class  $\alpha$  with  $e_\lambda$  the corresponding weight vector. By the discussion above, for  $U_{\alpha_0} = U_\alpha \circ \pi|_{G_0}$ ,  $\lambda|_{\mathfrak{t}_0} = \lambda_0 \in (\mathfrak{L}_0^*)_{K_0} \cap \bar{\mathfrak{C}}_0^*$  and  $e_\lambda = e_{\lambda_0}$  is a weight vector for  $U_{\alpha_0}$  corresponding to the highest weight  $\lambda_0$ . Let  $Q_{\alpha_1}$  be the projection to the subspace of vectors fixed by  $U_\alpha \circ \pi|_{K_1}$ . In view of (5) and subsection 2.1(2),

$$Q_{\alpha_1} e_\lambda = \frac{1}{r} \sum_{j=1}^r Q_{\alpha_0} U_\alpha(\exp_G H_j) e_\lambda = \frac{1}{r} \sum_{j=1}^r e^{2\pi i \lambda(H_j)} Q_{\alpha_0} e_{\lambda_0}.$$

However, by the definitions of  $\mathfrak{L}_K^*$  and  $K_1$ ,  $e^{2\pi i \lambda(H_j)} = 1$  for all  $j$  and hence  $Q_{\alpha_1} e_\lambda = Q_{\alpha_0} e_{\lambda_0} \neq 0$ . Hence  $Q_{\alpha_1} \neq 0$  and we conclude that  $\alpha \in \hat{G}_K$ .

Conversely, now suppose  $\alpha \in \hat{G}_K$  has highest weight  $\lambda(\alpha) = \lambda \in \mathfrak{C}^*$  and let  $\lambda_0 = \lambda|_{\mathfrak{t}_0}$ . For  $U_\alpha$ ,  $U_{\alpha_0}$ , and  $e_\lambda = e_{\lambda_0}$  as above, we know that  $\alpha_0 \in (\hat{G}_0)_{K_0}$  and  $Q_{\alpha_0} e_\lambda \neq 0$ . For any  $(G, K, \theta)$  satisfying (3), it is well known that the  $K$  bi-invariant functions in  $L^1(G, dx)$  commute under convolution [4, p. 408] and hence  $m_\alpha = 1$ . There is thus only a one dimensional span of vectors fixed by  $U_\alpha|_{K_1}$ ; any such vector is automatically fixed by  $U_{\alpha_0}$ . But the vectors fixed by  $U_{\alpha_0}|_{K_0}$  must also form a one-dimensional space and hence the spans coincide. For  $Q_{\alpha_1}$  as above, this simply means  $Q_{\alpha_1} = Q_{\alpha_0}$ . Now let  $H$  be any element in  $\mathfrak{t}$  for which  $\exp_G H \in K_1$ . Using (3) of 2.1 applied to  $Q_{\alpha_1}$  and  $K_1$ , we deduce that

$$0 \neq e^{2\pi i \lambda(H)} Q_{\alpha_0} e_\lambda = Q_{\alpha_0} (U_\alpha \circ \pi)(\exp_{G_1} H) e_\lambda = Q_{\alpha_0} e_\lambda$$

and hence  $\lambda(H) \in \mathbf{Z}$ . But then  $\lambda \in \mathfrak{L}_K^*$  and our proof is complete.

5.7. COROLLARY. Let  $M = K \setminus G$  be a compact Riemannian symmetric space. The Szegő theorem on  $M$ , i.e. (4) of §3, holds relative to the sequence  $\sigma(n)$  ( $n \geq 0$ ) defined using the notations of Theorem 5.6 by

$$\sigma(n) = \{ \alpha \in \hat{G}_K : \|\lambda(\alpha)\| \leq n \}$$

where  $\|\cdot\|$  is any Weyl group invariant norm on  $\mathfrak{t}^*$ .

PROOF. Define a polynomial  $f$  on the vector subspace of  $\mathfrak{t}^*$  spanned by  $\mathfrak{L}_K^*$  by

$$f(\lambda) = \prod_{\xi \in \Delta^+(\mathfrak{C})} \frac{\langle \lambda + \delta, \xi \rangle}{\langle \delta, \xi \rangle}.$$

By the discussion in subsection 5.2,  $f$  is positive on the intersection of this subspace with  $\bar{\mathfrak{C}}^*$ . As noted in the proof of Theorem 5.6,  $m_\alpha = 1$  for all  $\alpha \in \hat{G}_K$  and thus by Theorem 5.3(iii),

$$\sigma(n)^\# = \sum_{\alpha \in \sigma(n)} m_\alpha d_\alpha = \sum_{\substack{\lambda \in \mathfrak{L}_K^* \cap \bar{\mathfrak{C}}^* \\ \|\lambda\| \leq n}} f(\lambda).$$

From the proof of Theorem 5.4,  $\sigma_\gamma^\circ(n) \supset \sigma(n - \|\lambda(\gamma)\|)$  (see Lemma 3.3(v) for notations) for all  $\gamma \in \hat{G}_K$ . By applying Proposition 4.2, we conclude that condition (v) of Lemma 3.3 holds and thus Theorem 3.10 applies.

5.8. For  $k, m$  positive integers with  $k < m$ , the real Stiefel manifold  $S^{k,m}$  is defined as the set of all  $k \times m$  real matrices  $p$  for which  $p'p = I_k$ , the  $k \times k$  identity matrix. The group  $SO(m)$  of all real  $m \times m$  orthogonal matrices of determinant one

acts transitively on  $S^{k,m}$  from the right by ordinary matrix multiplication. Identifying  $SO(m - k)$  with the subgroup of  $SO(m)$  consisting of the  $m \times m$  matrices

$$\begin{bmatrix} I_k & 0 \\ 0 & u \end{bmatrix}$$

with  $u \in SO(m - k)$ , then  $SO(m - k)$  is the stability subgroup at the element  $p_0 = [I_k, 0] \in S^{k,m}$  and hence  $SO(m - k) \backslash SO(m)$  is a coset realization of  $S^{k,m}$ . For  $k = 1$ ,  $S^{1,m}$  is the unit sphere in  $\mathbf{R}^m$  and is a compact symmetric space. The discussion and results in 5.5–5.7 are not needed since it is well known that the representations of  $SO(m)$  appearing in  $L^2(S^{1,m})$  are just the usual spherical harmonic representations  $\alpha_i$  for  $i = 0, 1, 2, \dots$  with

$$d_{\alpha_i} = \binom{m + i - 1}{i} - \binom{m + i - 3}{i - 2}$$

and the spherical harmonic subrepresentations of  $\alpha_i \otimes \alpha_j$  are precisely those of the form  $\alpha_l$  with  $|i - j| \leq l \leq i + j$  and  $i + j - l$  even. Using this, one can easily construct many sequences  $\sigma(n)$  satisfying Lemma 3.3(v). For  $1 < k \leq m/2$ , Gelbart [2, p. 33] observed that with  $G = SO(m)$  and  $K = SO(m - k)$ , the highest weights of representations  $\alpha \in \hat{G}_K$  may be naturally identified with highest weights of finite dimensional irreducible representations of the noncompact group  $GL(k, \mathbf{R})$  in such a way that the multiplicity  $m_\alpha$  of  $\alpha$  in  $L^2(S^{k,m})$  is precisely the dimension of the corresponding representation of  $GL(k, \mathbf{R})$ . In particular,  $m_\alpha$  is given via Weyl's formula 5.3(iii) as a polynomial function of the highest weight of  $\alpha$ . This means that the quantities  $m_\alpha d_\alpha$  may again be viewed as the values of a polynomial function on lattice points in a cone. As in Theorem 5.4 and Corollary 5.7, Proposition 4.2 implies that the Szegő theorem is valid on  $M = S^{k,m}$  ( $k \leq \frac{m}{2}$ ) relative to the sequence  $\sigma(n)$  consisting of the elements in  $\hat{G}_K$  whose highest weight has length at most  $n$ . Similar results hold for the complex and quaternionic Stiefel manifolds  $U(m - k) \backslash U(m)$  and  $Sp(n - k) \backslash Sp(m)$  ( $k \leq n/2$ ).

For an arbitrary homogeneous space  $K \backslash G$  of a compact Lie group  $G$ , explicit description of  $m_\alpha$  by a closed form functional relationship is very difficult. In particular,  $m_\alpha$  is usually not described by a polynomial function of the highest weight of  $\alpha$ . The last author has computed  $m_\alpha$  explicitly for several low dimensional examples and has not found an example for which Lemma 3.3(v) cannot be satisfied for some  $\sigma$ ; indeed, in all of his examples, Lemma 3.3(v) is satisfied by the sequence used in Theorem 5.4. He therefore suspects that the Szegő theorem is valid on a much broader class of homogeneous manifolds than those examined here. It would be particularly interesting to prove its validity for  $K \backslash G$  with  $K$  a torus in a compact Lie group  $G$  since manifolds of this type arise in a variety of contexts. Here in low dimensional examples, it is possible to subdivide the cone containing all highest weights into smaller cones on each of which a polynomial estimate for  $m_\alpha d_\alpha$  is valid and then Lemma 3.3(v) is established via application of Proposition 4.2 to each of these smaller cones. We do not know whether the general case can be treated in this manner. Although there are presumably other families of homogeneous manifolds tractable by mild generalizations of the technique employed here, any fairly general

treatment will clearly require different methods. Outside the realm of compact Lie groups, groups  $G$  described as an infinite product of compact Lie groups can be easily handled in a straightforward way by using Theorem 5.4 on each factor to construct an appropriate net  $\sigma$ . Otherwise,  $\hat{G}$  usually does not permit an explicit description. It therefore appears that validity of the Szegő theorem for every compact  $G$  rests on whether an indirect argument can be given for the existence of a net  $\sigma$  satisfying any of the conditions in Lemma 3.3. Perhaps this might be done with the aid of Theorem 5.4 by viewing  $G$  as a projective limit of Lie groups.

#### REFERENCES

1. J. F. Adams, *Lectures on Lie groups*, Benjamin, New York, 1969.
2. S. Gelbart, *A theory of Stiefel harmonics*, Trans. Amer. Math. Soc. **192** (1974), 29–50.
3. A. Grenander and G. Szegő, *Toeplitz forms and their applications*, Univ. of California Press, Berkley, Calif., 1958.
4. S. Helgason, *Differential geometry and symmetric spaces*, Academic Press, New York, 1962.
5. E. Hewitt and K. Ross, *Abstract harmonic analysis*. II, Springer-Verlag, Berlin and New York, 1970.
6. J. E. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag, Berlin and New York, 1972.
7. H. R. Krieger, *Toeplitz operators on locally compact spaces*, J. Math. Mech. **14** (1965), 439–478.
8. David S. Liang, *Eigenvalue distributions of Toeplitz operators*, Thesis, Washington University, 1974.
9. F. Riesz and B. Sz. Nagy, *Functional analysis*, Ungar, New York, 1955.
10. G. Szegő, "Ein Grenzwertsatz über die Toeplitzischen Determinanten einer reel positiven Funktion", Math. Ann. **76** (1915), 490–503.
11. G. Warner, *Harmonic analysis on semi-simple Lie groups*. I, Springer-Verlag, Berlin and New York, 1972.

DEPARTMENT OF MATHEMATICS, WASHINGTON UNIVERSITY, ST. LOUIS, MISSOURI 63130 (Current address of I. I. Hirschman and E. N. Wilson)

Current address (D. S. Liang): Department of Mathematics, Benedict College, Columbia, South Carolina 29204