DECAY ON HOMOGENEOUS SPACES OF REDUCTIVE TYPE

BERNHARD KRÖTZ, EITAN SAYAG AND HENRIK SCHLICHTKRULL

Abstract. In this paper we explore homogeneous spaces $Z = G/H$ of a a real reductive Lie group G with a closed connected subgroup H. The investigation concerns the decay at infinity of smooth functions on Z, and L^p-integrability of matrix coefficients. These results are used in a study of the asymptotic density of lattice points on Z. Explicit examples are given of spaces for which results are new.

Date: June 6, 2011.
2000 *Mathematics Subject Classification.* 22F30, 22E46, 53C35, 22E40.

1. Introduction

The question of decay of matrix coefficients of unitary representations plays central role in the applications of harmonic analysis and representation theory to geometry, probability and number theory. In particular, the fundamental theorem of Howe-Moore, that matrix coefficients of infinite dimensional unitary representations decay to zero at infinity, has been an important ingredient in studying lattice counting on symmetric spaces [20], [22], [9].

The present paper establishes basic results concerning decay and integrability of generalized matrix coefficients, and more general smooth functions, in the context of unimodular homogeneous spaces.

The paper contains three types of results. We first establish a qualitative property of decay, valid for functions on general reductive homogeneous spaces. Next a more quantitative property of decay is established on spaces with refined structure and for matrix coefficients only. In particular, the decay allows one to conclude L^p-integrability. Finally the results on integrability, combined with the abstract Plancherel theorem, are used to study decay of periods of automorphic forms. Building on the work of [20], these results are applied to the asymptotic counting of lattice points on reductive homogeneous spaces.

1.1. **Qualitative results.** Let G be a real reductive group and $Z = G/H$ a unimodular homogeneous G space with a closed connected subgroup H. Qualitative bounds are developed for the decay of functions on Z. These bounds are expressed through the property *VAI* of Vanishing At Infinity. The space Z admits VAI if and only if for each $1 \leq p < \infty$, the smooth functions on Z, all of whose derivatives belong to $L^p(Z, \mu_Z)$, *vanish at infinity.*

Our main result in regards to this property is the following classification:

Theorem A. A unimodular homogeneous G-space $Z = G/H$ admits VAI if and only if $H \subset G$ is reductive.

The direction 'if' is proved in Proposition 4.7. If Z is of reductive type and $B \subset G$ is a compact ball we provide essentially sharp lower and upper bounds for $\mathrm{vol}_Z(Bz)$ where $z \in Z$ moves off to infinity (Section 4.2). These results generalize and simplify previous approaches in [36] and [31]. The lower bounds in particular imply that Z has VAI.

The converse implication is established in Proposition A.1 (see Appendix A). The main lemma shows that in the non-reductive case the volume of the above mentioned sets Bz can be made exponentially small.

1.2. **Quantitative results.** We impose on $Z = G/H$ the condition that there exists an open H orbit on a flag manifold G/P_{min} of G, with $P_{min} \subset G$ a minimal parabolic subgroup. Spaces that admit this property are said to be of *spherical type*.

Restricting analysis to spherical reductive homogenous spaces, we study the decay of matrix coefficients. More precisely, for a unitary irreducible representation π of G we consider the generalized matrix coefficient $m_{v,\eta}(g) := \langle v, \pi(g)\eta\rangle$ on G/H, with η an H-fixed distribution vector and v a smooth vector.

Assuming reductive spherical type, we obtain in Theorem 6.4 an upper bound for $m_{v,\eta}$ with a spectrally defined rate of decay. To formulate the result we denote by A the torus in the Iwasawa decomposition $G = KAN$, by A^+ its positive part and by Λ_π the exponent associated to π in [46] 4.3.5.

Theorem B. Suppose for the minimal parabolic subgroup $P \subset G$ that $\bar{P}H$ is open in G. Let V be a Harish-Chandra module for G and $\eta \in (V^{-\infty})^H$. Then there exists $d \in \mathbb{N}_0$ such that for all $v \in V$ there is a constant $C_v > 0$

$$|m_{v,\eta}(a)| \leq C_v a^{\Lambda_\pi}(1 + \|\log a\|)^d \qquad (a \in \overline{A^+})\,. \tag{1.1}$$

Theorem B is a generalization of [3], Thm. 6.1, and we view it as a basis for analysis on spherical spaces Z. We emphasize that Theorem B is applicable for general (non-unitary) Harish-Chandra modules. In the unitarizable case, the theorem of Howe-Moore implies that Λ_π is a sum of negative roots. Rather than generalizing the approach of [16] as was done in [3], we rest our arguments of [46] adapted to our situation. This yields a significantly shorter argument.

In many applications the L^p-integrability of the generalized matrix coefficients $m_{v,\eta}$ associated with unitary representations plays an important role. For example, the importance of matrix coefficient integrability was seen in [18], where the equivalence of an L^p-property with Kazhdan's Property (T) is shown.

In our case we express the desired integrability by a condition on Z, called *Property* (I), which asserts that for each irreducible unitary representation π of G there exists $p < \infty$ such that $m_{v,\eta} \in L^p(G/H_\eta)$ for all K-fixed vectors v and all H-fixed distribution vectors η. Here H_η is the stabilizer of η, often equal to H, but possibly strictly larger.

We believe that all reductive spherical type spaces admits this property (see Appendix D for our conjecture). The result we prove is weaker. To formulate it we introduce a class of reductive spaces, called *strong* spherical type (see definition 7.7). These spaces admit *polar* decomposition $G = KAH$ where K is maximal compact and A is a non-compact torus. Symmetric spaces are known to be of both polar and spherical type, and they satisfy the conditions of strong sphericality. It is our belief that all spaces of reductive spherical type are strongly spherical, a result which would make the stronger notion essentially superfluous. We provide a list of strongly spherical spaces below (the list is not exhaustive).

Our main result on integrability of generalized matrix coefficients of unitary representations is the following (see Theorem 8.5):

Theorem C. Suppose Z is of strong spherical type. Then Z admits Property (I).

The polar decomposition and the decay along A provided by Theorem B, yields the required integrability.

1.3. **Lattice counting.** In the last part of the paper we relate Property (I) to the problem of *lattice counting problems* which are situated at the crossroad of analysis, geometry and number theory.

Let $z_0 \in Z$ denote the base point. A *lattice* on G/H is the orbit $\Gamma.z_0$ of a lattice Γ in G for which $\Gamma_H = \Gamma \cap H$ is a lattice in H. The *lattice counting problem* on Z consists of the determination of the asymptotic behavior of the density of $\Gamma.z_0$ in K-invariant balls $B_R \subset Z$, as the radius $R \to \infty$. We note that the Gauss circle problem in the Euclidean plane is an instance of this problem and that Selberg developed his trace formula to establish lattice counting in the hyperbolic plane (see [39] for references). In the setting of semisimple symmetric spaces the problem was initiated by Duke, Rudnick and Sarnak in [44] and extensively studied in [20].

With proper normalizations of invariant measures, the *Main Term Counting* is the statement that the asymptotic density is 1. More precisely, with

$$N_R(\Gamma, Z) := \#\{\gamma \in \Gamma/\Gamma_H \mid \gamma.z_0 \in B_R\}$$

and $|B_R| := \mathrm{vol}(B_R)$ we have

$$N_R(\Gamma, Z) \sim |B_R| \quad (R \to \infty), \tag{1.2}$$

This was established in [20] for lattices on G/H for which H/Γ_H is compact. In subsequent work Eskin and McMullen [22] removed the obstruction and presented an ergodic approach. Later Eskin, Mozes and Shah [23] refined the ergodic methods and discovered that main term counting holds for a wider class of reductive spaces: It is valid whenever H is a maximal torus or a maximal reductive subgroup.

Our approach to the lattice counting problem is close in spirit to [20]. In proposition 10.2 we reformulate their approach in the form of a criterion. This criterion guarantees that main term counting holds as long as vanishing at infinity can be established for enough functions on Z, obtained as periods of automorphic functions on $\Gamma\backslash G$.

Based on (I), Plancherel decomposition and VAI we prove main term counting in Theorem 10.1 for symmetric spaces as well as for the spaces in the list below. We emphasize that in those cases H is not a maximal subgroup of G and main term counting is new.

Below we state a general result concerning main term counting which is applicable for any space admitting property (I). The statement, which is a simplification of 10.1 in the body of the paper, requires that the balls B_R *factorize well*, a simple notion introduced in Section 9.3 and related to the notion of *focusing* in [23].

These assumptions are verified for symmetric spaces and for an important class of geometric balls in all the examples listed below.

Theorem D. Let G be reductive and H a closed subgroup such that $Z = G/H$ is of reductive type. Suppose that $\Gamma \subset G$ is co-compact and Z admits (I). Suppose that $(B_R)_{R>0}$ factorizes well, then Main Term Counting holds.

Being of spectral nature, our approach extends and simplifies the method used in [20]. We believe that the co-compactness assumption can be removed, but this will require further results on regularization of periods of Eisenstein series that presently are not available in the literature (see comments in Appendix D).

In the end of Section 10 we consider a space for which Eskin and McMullen showed that main term counting fails to hold ([22]). It turns out that main term counting does hold for our family of geometric balls.

1.4. **Error Terms.** The problem of determining the error term in counting problems is notoriously difficult and in many cases relies on deep arithmetic information. Sometimes, like in the Gauss circle problem, some error term is easy to establish but getting an optimal error term is a very difficult problem.

We restrict ourselves to the cases where the cycle H/Γ_H is compact, and where Z is either symmetric or a *triple product space* as $\mathrm{SL}_2(\mathbb{R})^3/\Delta(\mathrm{SL}_2(\mathbb{R}))$, where Δ stands for the diagonal embedding.

The error we study is measure theoretic in nature, and will be denoted here as $\mathrm{err}(R, \Gamma)$. Thus, $\mathrm{err}(R, \Gamma)$ measures the deviation of two measures on $Y = \Gamma\backslash G$, the counting measure arising from lattice points in a ball of radius R, and the invariant measure μ_Y. It is easy to compare this error term with the pointwise error $\mathrm{err}_{pt}(R,\Gamma) = |N_R(\Gamma, Z) - |B_R||$, see Remark 11.1.

To formulate our result we introduce the exponent $p_H(\Gamma)$, which measures the worst L^p-behavior of any generalized matrix coefficient associated with a spherical unitary representation π, which is H-distinguished and occurs in the automorphic spectrum of $L^2(\Gamma\backslash G)$ (see definition 11.1). A related notion was studied in recent work of [30].

We first state our result for the non-symmetric case of triple product spaces, which is Theorem 12.4 from the body of the paper.

Theorem E. Let $Z = G_0^3/\operatorname{diag}(G_0)$ for $G_0 = \mathrm{SO}_e(1,n)$ and assume that H/Γ_H is compact. For all $p > p_H(\Gamma)$ there exists a $C = C(p) > 0$ such that

$$\mathrm{err}(R,\Gamma) \leq C|B_R|^{-\frac{1}{(6n+3)p}}$$

for all $R \geq 1$.

To the best of our knowledge this is the first error term obtained for a non-symmetric space. The crux of the proof is locally uniform comparison between L^p and L^∞ norms of generalized matrix coefficients which is achieved by applying the model of [12] for the triple product functional.

Let us now go back to the symmetric case. We will assume that $\Gamma_H\backslash H$ is compact. The existence of a non-quantitative error term for symmetric spaces was established in [9] and improved in [25].

Under these general assumptions we obtain error term estimates which in a considerable number of situations improve the bounds obtained in [25]. The next result is Theorem 11.8 in the body of the paper.

Theorem F. Let Z be symmetric. Assume

- $\Gamma_H = H \cap \Gamma$ is co-compact in H.
- $p > p_H(\Gamma)$
- $k > \frac{rank(G/K)+1}{2} \dim(G/K) + 1$

Then, there exists a constant $C = C(p,k) > 0$ such that

$$\mathrm{err}(R,\Gamma) \leq C|B_R|^{-\frac{1}{(2k+1)p}}$$

for all $R \geq 1$.

Moreover, if $Y = \Gamma\backslash G$ is compact one can replace the third condition by $k > \dim(G/K) + 1$.

The key estimate for the proof of Theorem F is Lemma 11.5 which is a quantitative version of property (I). To be more precise we compare in a uniform way the L^p and L^∞ norms. The proof found in Appendix C relies, among other things, on the Harish-Chandra-Gangolli recursion relations satisfied by spherical matrix coefficients.

We note that in case of the hyperbolic plane our error term is still far from the quality of the bound of A. Selberg. This is because we only use a week version of the trace formula, namely Weyl's law, and use simple soft Sobolev bounds between eigenfunctions on Y.

1.5. **List of examples.** To summarize, in this paper we prove that the following spaces are strongly spherical and hence, by Theorems C and D, admit Property (I) and Main Term Counting for the geometric balls.

(1) symmetric spaces
(2) some almost symmetric examples
   - $G = \mathrm{GL}(n+1, \mathbb{R})$ and $H = \operatorname{diag} \mathrm{GL}(n, \mathbb{R})$
   - $G = \mathrm{SL}(n+1, \mathbb{R})$ and $H = \mathrm{SL}(n, \mathbb{R})$
   - $G = \mathrm{Sp}(n+1, \mathbb{R})$ and $H = \mathrm{Sp}(n, \mathbb{R}) \times \mathbb{R}$
   - $G = \mathrm{U}(p, q+1, \mathbb{C})$ and $H = \mathrm{U}(p, q, \mathbb{C})$ with $p + q \geq 2$.
(3) Some G_2 related examples
   - $G = G_2(\mathbb{C})$ and $H = \operatorname{diag} \mathrm{SL}(3, \mathbb{C})$
   - $G = G_2(\mathbb{R})$ and $H = \operatorname{diag} \mathrm{SL}(3, \mathbb{R})$
   - $G = \mathrm{SO}(7, \mathbb{C})$ and $H = G_2(\mathbb{C})$
   - $G = \mathrm{SO}(8, \mathbb{C})$ and $H = G_2(\mathbb{C})$
(4) Gross-Prasad type spaces
   - $G = \mathrm{GL}(n+1, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R})$ and $H = \operatorname{diag} \mathrm{GL}(n, \mathbb{R})$
   - $G = \mathrm{U}(p, q+1, \mathbb{F}) \times \mathrm{U}(p, q, \mathbb{F})$ and $H = \operatorname{diag} \mathrm{U}(p, q, \mathbb{F})$

   Here $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.
(5) Triple product spaces
   - $G = G_0 \times G_0 \times G_0$ and $H = \operatorname{diag} G_0$

   where $G_0 = \mathrm{SO}_e(1, n)$ and $n \geq 2$.
(6) Some non-maximal subgroups that are non-symmetric.
   - $G = \mathrm{SL}(2n+1, \mathbb{R})$ and $H = \mathrm{Sp}(n, \mathbb{R})$
   - $G = \mathrm{SU}(n, n+1)$ and $H = \mathrm{Sp}(k, k)$ with $n = 2k$.
   - $G = \mathrm{SO}(n, n+1)$ and $H = \mathrm{U}(k, k)$ with $n = 2k$.
   - $G = \mathrm{SU}(n, n+1)$ and $H = \mathrm{Sp}(k, k+1)$ with $n = 2k+1$.
   - $G = \mathrm{SO}(n, n+1)$ and $H = \mathrm{U}(k, k+1)$ with $n = 2k+1$.

Acknowledgment: The authors would like to thank Joseph Bernstein, Heiko Gimperlein and Erez Lapid for several enlightening discussions on various aspects of the paper.

2. Vanishing at infinity

Let G be a real Lie group and $H \subset G$ a closed subgroup. Consider the homogenous space $Z = G/H$ and assume that it is unimodular, that is, it carries a G-invariant measure μ_Z. Note that such a measure is unique up to a scalar multiple.

For a Banach representation (π, E) of G let us denote by E^∞ the space of smooth vectors. In the special case for the left regular representation of G on $E = L^p(Z)$ with $1 \leq p < \infty$, it follows from the local Sobolev lemma that $E^\infty \subset C^\infty(Z)$. Let $C_0^\infty(Z)$ be the space of smooth functions on Z that vanish at infinity. Motivated by the decay of eigenfunctions on symmetric spaces ([43]), the following definition was taken in [34]:

Definition 2.1. *We say Z has the property* VAI *(vanishing at infinity) if for all $1 \leq p < \infty$ we have*

$$L^p(Z)^\infty \subset C_0^\infty(Z).$$

By a result of [40], $Z = G$ has the VAI property for G unimodular and $H = \{\mathbf{1}\}$. The main result of [34] establishes that all reductive symmetric spaces admit VAI. On the other hand, if H is a non-cocompact lattice in G then $Z = G/H$ is not VAI.

Let G be a real reductive group (see [46]). We say that H is a reductive subgroup of G and that Z is *of reductive type* (or just reductive), if H is real reductive and the adjoint representation of H in the Lie algebra $\mathfrak{g}$ of G is completely reducible. Note that Z is unimodular in this case.

Theorem 2.2. *Let G be a real reductive group and $H \subset G$ a closed connected subgroup such that $Z = G/H$ is unimodular. Then* VAI *holds for Z if and only if it is of reductive type.*

3. The invariant measure

In this section we provide a suitable framework for a discussion of the invariant measure on Z. Throughout G is a connected real reductive group and $H \subset G$ is a closed connected subgroup such that $Z := G/H$ is unimodular.

Let $\mathfrak{g}$ and $\mathfrak{h}$ be the Lie algebras of G and H. We fix a Levi-decomposition $\mathfrak{h} = \mathfrak{r} \rtimes \mathfrak{s}$ of $\mathfrak{h}$ and choose a subalgebra $\mathfrak{l} \supset \mathfrak{s}$ which is reductive in $\mathfrak{g}$ and contained in $\mathfrak{h}$, and which is maximal with these properties. Then $Z = G/H$ is of reductive type if and only if $\mathfrak{h} = \mathfrak{l}$. Let L be the analytic subgroup of H corresponding to the subalgebra $\mathfrak{l}$.

Using [38], Ch. 6, Thm. 3.6, we fix a Cartan involution of G which preserves L, and such that the restriction to L is a Cartan involution of L. The derived involution $\mathfrak{g} \to \mathfrak{g}$ will also be called θ. We may and shall assume that $\theta(\mathfrak{s}) = \mathfrak{s}$.

The fixed point set of θ determines a maximal compact subgroup K of G whose Lie algebra will be denoted $\mathfrak{k}$. Let $\mathfrak{p}$ denote the -1-eigenspace of θ on $\mathfrak{g}$, then $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Let κ be a non-degenerate invariant symmetric bilinear form on $\mathfrak{g}$ such that

- $\kappa|_{\mathfrak{p}} > 0$
- $\kappa|_{\mathfrak{k}} < 0$
- $\mathfrak{k} \perp \mathfrak{p}$ with respect to κ.

Having defined κ we define an inner product on $\mathfrak{g}$ by $\langle X, Y\rangle = -\kappa(\theta(X), Y)$. Let $\mathfrak{q}$ be the orthogonal complement of $\mathfrak{h}$ in $\mathfrak{g}$.

Remark 3.1. Let Z be of reductive type. Then $[\mathfrak{h}, \mathfrak{q}] \subset \mathfrak{q}$. Moreover one has $[\mathfrak{q}, \mathfrak{q}] \subset \mathfrak{h}$ if and only if the pair $(\mathfrak{g}, \mathfrak{h})$ is symmetric, that is, if and only if $\mathfrak{h} = \{X \in \mathfrak{g} \mid \sigma(X) = X\}$ for an involution σ of $\mathfrak{g}$. Then $\mathfrak{q} = \{X \in \mathfrak{g} \mid \sigma(X) = -X\}$

3.1. **Construction of the invariant measure.** The differential geometric way to obtain an invariant measure on Z is by defining an invariant differential form of top degree. Let us briefly recall this construction.

For every $g \in G$ we denote by

$$\tau_g : Z \to Z, \quad xH \mapsto gxH$$

the diffeomorphic left displacement by g on Z. Let $z_0 = H \in Z$ be the base point. Given $g \in G$ we shall identify the tangent space $T_{gz_0}Z$ of Z at the point gz_0 with $\mathfrak{g}/\mathfrak{h}$ via the map

$$\mathfrak{g}/\mathfrak{h} \to T_{gz_0}Z, \quad X + \mathfrak{h} \mapsto d\tau_g(z_0)X\,. \tag{3.1}$$

Let us emphasize that if $gz_0 = g'z_0$, then $g = g'h$ for some $h \in H$ and the two identifications differ by the automorphism $\mathrm{Ad}(h)|_{\mathfrak{g}/\mathfrak{h}}$. The assumption that an invariant measure exists on Z implies that the determinant of this automorphism is 1.

Let $Y_1, ..., Y_s$ be a basis of $\mathfrak{g}/\mathfrak{h}$ and $\omega_1, ..., \omega_s$ the corresponding dual basis in $(\mathfrak{g}/\mathfrak{h})^* \subset \mathfrak{g}^*$. We define the H-invariant volume form on $\mathfrak{g}/\mathfrak{h}$ by

$$\omega = \omega_1 \wedge \omega_2 \wedge ... \wedge \omega_s \in \bigwedge^s (\mathfrak{g}/\mathfrak{h})^*\,.$$

As ω is $\mathrm{Ad}(H)$-invariant we can extend ω to a G-invariant volume form ω_Z on Z. The measure μ_Z corresponding to ω_Z is then a Haar measure on Z.

4. Reductive Spaces are VAI

In this section $Z = G/H$ is of reductive type. Our goal is to establish uniform bounds for the invariant measure and deduce VAI for these spaces. As Z is of reductive type we can and will identify $\mathfrak{g}/\mathfrak{h}$ with $\mathfrak{q}$ in an H-equivariant way. Note that $\mathfrak{q}$ is θ-stable and in particular $\mathfrak{q} = \mathfrak{q} \cap \mathfrak{k} + \mathfrak{q} \cap \mathfrak{p}$. We denote by $\mathrm{pr}_{\mathfrak{q}} : \mathfrak{g} \to \mathfrak{q}$ the orthogonal projection.

The characterization of infinity on Z is obtained by the polar decomposition which asserts that the polar map

$$\pi : K \times_{H\cap K} (\mathfrak{q} \cap \mathfrak{p}) \to Z, \quad [k, Y] \mapsto k \exp(Y) z_0 \tag{4.1}$$

is a homeomorphism (see [37]). Then a function $f \in C(Z)$ vanishes at infinity if and only if

$$\lim_{\substack{Y \mapsto \infty \\ Y \in \mathfrak{q}\cap\mathfrak{p}}} \sup_{k\in K} |f(\pi(k, Y))| = 0\,.$$

4.1. Local coordinates.

The objective of this subsection is to provide some useful local coordinates on Z and to give a uniform estimate of the invariant measure in terms of these local coordinates.

Let $U_R = \{X \in \mathfrak{g} : ||X|| < R\}$ for $R > 0$ and $U_{R,\mathfrak{q}} = U_R \cap \mathfrak{q}$. Note that when R is sufficiently small then $\exp|_{U_R}$ is diffeomorphic onto its image in G. We define for all $R > 0$ a 'ball' in G by $B_{R,\mathfrak{g}} = \exp(U_R)$. Likewise we define $B_{R,\mathfrak{q}} = \exp(U_{R,\mathfrak{q}}) \subset G$.

Let $g \in G$ and define a map ϕ_g by

$$\phi_g : U_{R,\mathfrak{q}} \to Z, \quad Y \mapsto \exp(Y) g z_0\,.$$

Observe that

$$\mathrm{vol}_Z(B_{R,\mathfrak{g}} g z_0) \geq \mathrm{vol}_Z(B_{R,\mathfrak{q}} g z_0) = \int_{U_{R,\mathfrak{q}}} \phi_g^* \omega_Z$$

with the last equality holding if ϕ_g is diffeomorphic onto its image.

We will now show that ϕ_g is a coordinate chart with a Jacobian uniformly bounded from below provided $g = \exp(X)$ with $X \in \mathfrak{p} \cap \mathfrak{q}$ sufficiently large. We shall identify $T_{\exp(Y)gz_0} Z$ with $\mathfrak{q}$ as in (3.1). A standard computation yields for all $Y \in U_{\mathfrak{q},R}$:

$$\begin{aligned} d\phi_g(Y) : X \mapsto \mathrm{pr}_{\mathfrak{q}}(\mathrm{Ad}(g^{-1}) \left(\frac{1 - e^{-\operatorname{ad} Y}}{\operatorname{ad} Y} X \right)), \\ \mathfrak{q} \to T_{\exp(Y)gz_0} Z = \mathfrak{q} \end{aligned} \tag{4.2}$$

For $Y \in U_{\mathfrak{q},R}$ we shall denote by

$$J_g(Y) = |\det d\phi_g(Y)|$$

the Jacobian of ϕ_g at Y.

Lemma 4.1. *There exists a neighborhood $U \subset \mathfrak{q}$ of 0 and constants $C, d > 0$ such that*

$$J_{\exp(X)}(Y) \geq d \tag{4.3}$$

for all $X \in \mathfrak{p} \cap \mathfrak{q}$ with $\|X\| \geq C$, and all $Y \in U$. In particular, the map

$$\phi_{\exp X} : U \to Z$$

is then diffeomorphic onto its image.

Proof. We may assume that the basis $Y_1, \dots, Y_s$ of $\mathfrak{g}/\mathfrak{h}$ is an orthonormal basis of $\mathfrak{q}$. It then follows from (4.2) that

$$J_{\exp X}(Y) = |\det\left(\langle e^{-\operatorname{ad} X} \circ \left(\frac{1 - e^{-\operatorname{ad} Y}}{\operatorname{ad} Y}\right) Y_i, Y_j\rangle_{1 \leq i,j \leq s}\right)|. \tag{4.4}$$

Since $\theta(X) = -X$ we can rewrite the matrix elements in (4.4) as

$$\langle\left(\frac{1 - e^{-\operatorname{ad} Y}}{\operatorname{ad} Y}\right) Y_i, e^{-\operatorname{ad} X} Y_j\rangle\,. \tag{4.5}$$

Observe that $\operatorname{ad} X$ is real semisimple and let $V_1, \dots, V_n$ be an orthonormal basis for $\mathfrak{g}$ of eigenvectors, with corresponding real eigenvalues $\lambda_1, ..., \lambda_n$. Let

$$b_{ik} = \langle\left(\frac{1 - e^{-\operatorname{ad} Y}}{\operatorname{ad} Y}\right) Y_i, V_k\rangle\,, \qquad c_{kj} = \langle V_k, Y_j\rangle,$$

then (4.5) equals $\sum_{k=1}^n b_{ik} c_{kj} e^{-\lambda_k}$. The determinant in (4.4) is a sum of products of such expressions.

We replace X by tX for $t \in \mathbb{R}$ and set

$$p(t) = p_{X,Y}(t) := \det\left(\langle\left(\frac{1 - e^{-\operatorname{ad} Y}}{\operatorname{ad} Y}\right) Y_i, e^{-\operatorname{ad} tX} Y_j\rangle_{1 \leq i,j \leq s}\right).$$

Then it follows from the reasoning above that p is a linear combination of exponential functions $e^{-\lambda t}$ with exponents $\lambda \in \mathbb{R}$ which are sums of eigenvalues $-\lambda_k$. We observe that the exponents depend on X in a way which can be arranged to be locally uniform, and likewise for the dependence of the coefficients on X and Y.

For $Y = 0$ we have $p_{X,0}(-t) = p_{X,0}(t)$ for all t and all X, since $\theta(X) = -X$ and $\mathfrak{q}$ is θ-invariant. Thus $e^{\lambda t}$ and $e^{-\lambda t}$ will occur with the same coefficients in the expansion of $p_{X,0}$. If we denote by λ_X the maximal exponent λ such that $e^{\lambda t}$ occurs in $p_{X,0}(t)$ with a non-zero coefficient, then we conclude that $\lambda_X \geq 0$. By compactness and local uniformity it follows that there exists a compact neighborhood $U \subset \mathfrak{q}$ of 0 such that $e^{\lambda_X t}$ occurs in the expansion of $p_{X,Y}(t)$ with non-zero coefficient for all $Y \in U$ and any unit vector $X \in \mathfrak{p} \cap \mathfrak{q}$.

In the expansion of $p(t)$ the term with maximal exponent λ will dominate the others when $t \to \infty$. As we have just seen, this maximal exponent is ≥ 0 for all $Y \in U$ and all X. Hence there exist constants $C, d > 0$ such that $|p(t)| \geq d$ for $t > C$. Again by compactness, these constants can be chosen independently of $Y \in U$ and X with $\|X\| = 1$. ☐

In general, the constant lower bound (4.3) is sharp. However, in many cases one can improve to an exponential lower bound. A particularly simple case is obtained when $(\mathfrak{g}, \mathfrak{h})$ is a symmetric pair. Let $\mathrm{pr}_s : \mathfrak{q} \to \mathfrak{q} \cap [\mathfrak{g}, \mathfrak{g}]$ denote the orthogonal projection of $\mathfrak{q}$ to its semisimple part, and let $X_s = \mathrm{pr}_s(X)$ for $X \in \mathfrak{q}$.

Lemma 4.2. *Assume G/H is a symmetric space. Then there exists a neighborhood $U \subset \mathfrak{q}$ of 0 and constants $C, d, \delta > 0$ such that*

$$J_{\exp(X)}(Y) \geq de^{\delta\|X_s\|}$$

for all $X \in \mathfrak{p} \cap \mathfrak{q}$ with $\|X\| \geq C$, and all $Y \in U$.

Proof. We denote by σ the involution of $\mathfrak{g}$ associated with $\mathfrak{h}$. Let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p}$ containing X and let Σ be the associated system of restricted roots. Root spaces in $\mathfrak{g}$ are denoted $\mathfrak{g}^\alpha$, where $\alpha \in \Sigma$. Let

$$\delta_X = \sum_{\alpha \in \Sigma, \alpha(X) > 0} \alpha(X)$$

(roots counted with multiplicities), then δ_X is independent of the choice of subspace $\mathfrak{a}$, and $\delta_X \geq \delta\|X_s\|$ for some constant $\delta > 0$, independent of X. We claim that there exist U, C and d as above such that

$$J_{\exp(X)}(Y) \geq de^{\delta_X} \tag{4.6}$$

for all $Y \in U$ and $\|X\| \geq C$. Obviously this will imply the lemma.

Notice that the $\mathrm{ad}\, X$-eigenspace for the eigenvalue $\lambda \in \mathbb{R}$ is given by $\mathfrak{g}_X^\lambda = \oplus_{\alpha \in \Sigma, \alpha(X) = \lambda} \mathfrak{g}^\alpha$. Note also that $V \in \mathfrak{g}_X^\lambda$ implies $\sigma(V) \in \mathfrak{g}_X^{-\lambda}$, since $\sigma(X) = -X$.

We follow the proof of Lemma 4.1. It suffices to prove that if λ_X is the maximal exponent in this proof then $\lambda_X t = \delta_{tX}$ for $t > 0$. It follows from the preceding paragraph that we can choose the orthonormal basis $(V_k)_{1,\dots,n}$ for $\mathfrak{g}$ such that basis vectors with non-zero eigenvalues $\pm\lambda$ for $\mathrm{ad}\, X$ are mutually paired by σ, and that each root with $\alpha(X) > 0$ is represented by such pairs according to its multiplicity. The orthonormal basis $(Y_j)_{1,\dots,s}$ for $\mathfrak{q}$ can then be chosen to consist of normalized multiples of the $V_k - \sigma(V_k)$ for each such pair, and additional vectors commuting with X. An elementary computation now shows that

$$p_{X,0}(t) = \Pi_{\alpha \in \Sigma, \alpha(X) > 0} \cosh(\alpha(tX)),$$

from which the expression for the maximal exponent follows. □

4.2. **Volume bounds.** We record the following corollaries.

Corollary 4.3. *There exist a neighborhood $U \subset \mathfrak{q}$ of 0 and constants $C, d > 0$ such that the following holds for all $X \in \mathfrak{q} \cap \mathfrak{p}$ with $\|X\| > C$: For each $R > 0$ with $U_{R,\mathfrak{q}} \subset U$,*

$$\mathrm{vol}_Z(B_{R,\mathfrak{q}} \exp(X) z_0) \geq d \,\mathrm{vol}_{\mathfrak{q}}(U_{R,\mathfrak{q}}).$$

When G/H is symmetric, the lower volume bound can be improved to $de^{\delta\|X_s\|}\,\mathrm{vol}_{\mathfrak{q}}(U_{R,\mathfrak{q}})$ with $\delta > 0$ independent of X and R.

Remark 4.4. The proof of Lemma 4.1 also provides the following upper volume bound. There exist constants $D, \lambda > 0$ such that

$$\mathrm{vol}_Z(B_{R,\mathfrak{q}} \exp(X) z_0) \leq De^{\lambda\|X\|}\,\mathrm{vol}_{\mathfrak{q}}(U_{R,\mathfrak{q}})$$

for all $X \in \mathfrak{q} \cap \mathfrak{p}$. See [31] for such a bound in the literature.

Corollary 4.5. *There exist constants $C, d, R_0 > 0$ such that*

$$\mathrm{vol}_Z(B_{R,\mathfrak{g}} z) \geq dR^{\dim \mathfrak{q}}$$

for all $R \leq R_0$ and all $z = k \exp(X) z_0 \in Z$ with $k \in K$, $X \in \mathfrak{p} \cap \mathfrak{q}$ and $\|X\| > C$. Furthermore, in the symmetric case we have

$$\mathrm{vol}_Z(B_{R,\mathfrak{g}} z) \geq de^{\delta\|X_s\|} R^{\dim \mathfrak{q}}$$

with $\delta > 0$ independent of R and z.

Remark 4.6. For fixed $R > 0$ and G, H semisimple it was shown in [36] that there exists a constant $c > 0$ such that

$$\mathrm{vol}_Z(B_{R,\mathfrak{g}} z) > c$$

for all $z \in Z$. The corollary above sharpens this bound.

4.3. **Vanishing at infinity.**

Proposition 4.7. *Let Z be a homogeneous space of reductive type. Then Z has the property* VAI. *Moreover, the inclusion*

$$L^p(Z)^\infty \subset C_0^\infty(Z)$$

is continuous.

Proof. By applying the Sobolev inequality in local coordinates, we obtain the following for $1 \leq p < \infty$ and for each compact neighborhood B of e in G (see [40] for details). There exist finitely many elements $v_i \in \mathcal{U}(\mathfrak{g})$ (of degree up to the smallest integer $> \dim Z/p$) in the enveloping algebra of $\mathfrak{g}$, and for each $z \in Z$ a constant $D > 0$ such that

$$|f(z)| \leq D \max \|(L_{v_i} f)\, \mathbf{1}_{Bz}\|_p \tag{4.7}$$

for all $f \in L^p(Z)^\infty$. Here $\mathbf{1}_{Bz}$ denotes the characteristic function of $Bz \subset Z$. The constant D is locally uniform with respect to z.

Based on Lemma 4.1, we can improve the local estimate (4.7), such that for G/H of reductive type it holds with D independent of z. Let $U \subset \mathfrak{q}$ and $C, d > 0$ be as in Lemma 4.1, and fix $R > 0$ such that $U_{R,\mathfrak{q}} \subset U$. It follows that

$$\|(f \circ \phi_{\exp X})\mathbf{1}_{U_{R,\mathfrak{q}}}\|_p \leq d^{-1/p}\|f\, \mathbf{1}_{B_{R,\mathfrak{q}} \exp(X) z_0}\|_p \tag{4.8}$$

for $f \in L^p(Z)$ and $X \in \mathfrak{p} \cap \mathfrak{q}$ with $\|X\| > C$. By the Sobolev inequality for $\mathbb{R}^{\dim \mathfrak{q}}$, the value $|f \circ \phi_{\exp X}(0)|$ is estimated above by the p-norms, over any neighborhood of 0, of the derivatives of $f \circ \phi_{\exp X}$. Hence if $f \in L^p(Z)^\infty$ and $\|X\| > C$, we obtain an upper bound

$$|f(\exp(X)z_0)| \leq D \max \|(L_v f)\, \mathbf{1}_{B_{R,\mathfrak{q}} \exp(X) z_0}\|_p$$

with derivatives as before by finitely many elements in $\mathcal{U}(\mathfrak{g})$, and with a constant D independent of f and X. After conjugation by $k \in K$ we conclude that (4.7) holds at $z = k \exp(X) z_0$, with $B = B_{R,\mathfrak{g}}$ and with a uniform constant D. As the set of elements $k \exp(X) z_0$ with $\|X\| \leq C$ is compact, the inequality is finally obtained for all $z \in Z$.

The proposition is a straightforward consequence of the uniform version of (4.7). □

For G/H symmetric we obtain a better result by replacing the use of Lemma 4.1 by Lemma 4.2 in the estimate (4.8) while following the preceding proof:

Proposition 4.8. *Let $Z = G/H$ be symmetric. There exists a constant $\delta > 0$ with the following property. Let $f \in L^p(Z)^\infty$ where $1 \leq p < \infty$. Then for each $\epsilon > 0$ there exists $C > 0$ such that*

$$|f(k \exp(X) z_0)| \leq \epsilon e^{-\delta \|X_s\|}$$

for all $z = k \exp(X) z_0$, where $X \in \mathfrak{p} \cap \mathfrak{q}$, $\|X\| > C$ and $k \in K$.

5. Homogeneous spaces of polar type

Recall that all homogeneous spaces of reductive type admit the decomposition (4.1). A stronger decomposition property can be defined as follows.

Definition 5.1. *Let $Z = G/H$ be a homogeneous space of reductive type. A* polar decomposition *of Z consists of an abelian subspace $\mathfrak{a} \subset \mathfrak{p}$ and a surjective proper map*

$$K \times A \ni (k, a) \mapsto kaH \in Z,$$

where $A = \exp \mathfrak{a}$. *We say that* Z *is of* polar type *if such a polar decomposition exists.*

Notice that we do not require $\mathfrak{a} \subset \mathfrak{q}$ in Definition 5.1. This is not possible for example in the spaces considered in 5.1.2 and 5.1.3.

5.1. **Examples.** We provide some examples of homogeneous spaces of polar type. Further examples will be given later (see Section 6.1)

5.1.1. *Symmetric spaces.* Symmetric spaces are of polar type with $\mathfrak{a} \subset \mathfrak{q}$. In fact (see [45] p. 117) suppose that Z is symmetric and let $\mathfrak{a}_q \subset \mathfrak{p} \cap \mathfrak{q}$ be a maximal abelian subspace (it is unique up to conjugation by $K \cap H$). Then $G = KA_qH$ where $A_q = \exp(\mathfrak{a}_q)$. For this case we recall that more structure is known. Associated to $\mathfrak{a}_q$ is a root system $\Sigma(\mathfrak{g}, \mathfrak{a}_q) \subset \mathfrak{a}_q^*$ and hence a notion of positivity. For each associated Weyl chamber $\mathfrak{a}_q^+$ the map

$$K \times_M \mathfrak{a}_q^+ \to Z, \quad [k, X] \mapsto k \exp(X) z_0$$

is a diffeomorphism onto an open set. Here $M = Z_{K \cap H}(\mathfrak{a}_q)$. Furthermore, the union of the sets $KA_q^+ z_0$, over all $K \cap H$-conjugacy classes of positive chambers $A_q^+ = \exp(\mathfrak{a}_q^+)$, is disjoint and dense in Z.

Remark 5.2. Although not subject proper of this paper it is quite useful to compare the situation in the p-adic setup: Let G be a p-adic reductive group G, $K < G$ a maximal compact subgroup and $H < G$ a symmetric subgroup. Then, in general, there exists no torus A such that $G = KAH$. However, there exists a torus $A < G$ and a finite subset F such that $G = KAFH$, see [19].

5.1.2. *Triple spaces.* Let G_0 be a real reductive group and define

$$G = G_0^3 = G_0 \times G_0 \times G_0$$

and $H = \operatorname{diag}(G_0) = \{(g, g, g) \mid g \in G_0\}$. Then $Z = G/H$ is a homogeneous space of reductive type.

Let $K_0 < G_0$ be a maximal compact subgroup. We fix an Iwasawa decomposition $G_0 = K_0 A_0 N_0$ and let $P_0 = M_0 A_0 N_0$ be the associated minimal parabolic subgroup. Set $K = K_0 \times K_0 \times K_0$.

Proposition 5.3. *Suppose that* $B_0 \subset G_0$ *is a subset such that*

$$G_0 = A_0 M_0 B_0 K_0.$$

Then, for $A = A_0 \times A_0 \times B_0$, *one has* $G = KAH$.

Proof. Let $(g_1, g_2, g_3) \in G$. From the KAH-decomposition of the symmetric space $G_0 \times G_0 / \operatorname{diag}(G_0)$ we obtain

$$(g_1, g_2) = (g, g)(a_1, a_2)(k_1, k_2)$$

for some $g \in G_0$, $a_1, a_2 \in A_0$ and $k_1, k_2 \in K_0$. Now choose $m \in M_0$, $a_0 \in A_0$, $b_0 \in B_0$ and $k_0 \in K_0$ such that $g^{-1}g_3 = a_0 m_0 b_0 k_0$. Then

$$(g_1, g_2, g_3) = (ga_0m_0, ga_0m_0, ga_0m_0)(a_0^{-1}a_1, a_0^{-1}a_2, b_0)(m_0^{-1}k_1, m_0^{-1}k_2, k_0)$$

as asserted. □

The proposition applies in the following cases:

- $G_0 = \mathrm{SO}_e(1,n)$ and $B_0 = \exp(\mathbb{R}X)$ for some $0 \neq X \in \mathfrak{p}_0 \cap \mathfrak{a}_0^\perp$.

Note that it also applies to $B_0 = N_0$ for general G_0,

Corollary 5.4. *Let* $G_0 = \mathrm{SO}_e(1,n)$ *for* $n \geq 2$ *and* $Z = G_0^3/\operatorname{diag}(G_0)$. *Then* Z *is of polar type.*

5.1.3. *Gross-Prasad spaces.* We let G_0 be a reductive group and $H_0 < G_0$ be a reductive subgroup. Set $G = G_0 \times H_0$ and $H = \operatorname{diag}(H_0)$. Note that $Z = G/H \simeq G_0$. However in this isomorphism G_0 is viewed as a homogeneous $G_0 \times H_0$ space.

We consider the following choices for G_0 and H_0, with which we refer to Z as a Gross-Prasad space (cf. [26]):

- $G_0 = \mathrm{GL}(n+1,\mathbb{R})$ and $H_0 = \mathrm{GL}(n,\mathbb{R})$ for $n \geq 0$.
- $G_0 = \mathrm{U}(p, q+1, \mathbb{F})$ and $H_0 = \mathrm{U}(p,q,\mathbb{F})$ for $p+q \geq 2$.

Here $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.

Lemma 5.5. *Gross-Prasad spaces are of polar type.*

Proof. We first treat the case $(G_0, H_0) = (\mathrm{GL}(n+1,\mathbb{R}), \mathrm{GL}(n,\mathbb{R}))$. Let us embed H_0 in G_0 as the lower right corner.

We define a two-dimensional non-compact torus of $\mathrm{GL}(2,\mathbb{R})$ by

$$B = \mathrm{SO}_e(1,1) \cdot \mathbb{R}_+\mathbf{1}\,.$$

In $\mathrm{GL}(2k,\mathbb{R})$ we define a $2k$-dimensional non-compact torus A_{2k} by k block matrices of form B along the diagonal. In $\mathrm{GL}(2k+1,\mathbb{R})$ we define A_{2k+1} to consist of similar blocks together with a positive number in the last diagonal entry. Finally we let

$$A = A_{n+1} \times A_n \subset G = \mathrm{GL}(n+1,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R}).$$

With $K = \mathrm{O}(n+1,\mathbb{R}) \times \mathrm{O}(n,\mathbb{R})$ we claim that

$$G = KAH\,,$$

or, equivalently,

$$\mathrm{GL}(n+1,\mathbb{R}) = \mathrm{O}(n+1,\mathbb{R}) A_{n+1} A_n \, \mathrm{O}(n,\mathbb{R})\,.$$

We proceed by induction on n. The case $n = 0$ is clear. We shall use the known polar decomposition for the almost symmetric pair $(\mathrm{GL}(n+1, \mathbb{R}), \mathrm{GL}(n, \mathbb{R}))$:

$$\mathrm{GL}(n+1, \mathbb{R}) = \mathrm{O}(n+1, \mathbb{R}) B_1 \, \mathrm{GL}(n, \mathbb{R})$$

where B_1 is the two-dimensional torus of form B located in in the upper left corner. Now insert for $\mathrm{GL}(n, \mathbb{R})$ by induction, but in opposite order:

$$\mathrm{GL}(n, \mathbb{R}) = \mathrm{O}(n-1, \mathbb{R}) A_{n-1} A_n \, \mathrm{O}(n, \mathbb{R})$$

and observe that $\mathrm{O}(n-1, \mathbb{R})$ commutes with B_1.

The case with $G_0 = \mathrm{U}(p, q+1, \mathbb{F})$ is similar. Choose non-compact Cartan subspaces for $\mathfrak{g}_0$ and $\mathfrak{h}_0$ along antidiagonals, and note that the overlap between these, as subspaces of $\mathfrak{g}_0$, is trivial. Now proceed by induction as before. □

Remark 5.6. Note that in Corollary 5.4 and Lemma 5.5 all polar decompositions $G = KAH$ are with $\mathfrak{a}$ a Cartan subspace in $\mathfrak{p}$.

5.2. **Some structure theory.** Let $Z = G/H$ be a homogeneous space of reductive type. We assume that the Cartan involution θ is chosen such that $\theta(H) = H$.

Lemma 5.7. *There exists a finite dimensional representation* (π, V) *of* G *and a vector* $v_{\mathfrak{h}} \in V$ *such that* $\mathfrak{h} = \{X \in \mathfrak{g} | d\pi(X) v_{\mathfrak{h}} = 0\}$.

Proof. Follows from Sect. 5.6, Th. 3 in [2]. □

Let $\mathfrak{a} \subset \mathfrak{p}$ be an abelian subspace with $\mathfrak{a} \cap \mathfrak{h} = \{0\}$. Let $A = \exp(\mathfrak{a})$.

Lemma 5.8. *The set* AH *is closed in* G *and* $(a, h) \mapsto ah$ *is proper* $A \times H \to AH$.

Proof. We argue by contradiction. Suppose that AH were not closed in G or that $(a, h) \mapsto ah$ were not proper. Then there would exist sequences $(a_n)_{n \in \mathbb{N}}$ in A and $(h_n)_{n \in N}$ in H, both leaving every compact subset, such that $p = \lim_{n \to \infty} a_n h_n$ exists in G.

Let (π, V) be a finite dimensional representation as in Lemma 5.7. Then the limit $\lim_{n \to \infty} \pi(a_n) v_{\mathfrak{h}}$ exists. Passing to a subsequence we may assume that

$$X := \lim_{n \to \infty} \frac{\log a_n}{\| \log a_n \|} \in \mathfrak{a}$$

exists. This implies that $v_{\mathfrak{h}}$ is a sum of $d\pi(X)$-eigenvectors with non-positive eigenvalues. Likewise $(\theta(a_n h_n))_{n \in \mathbb{N}}$ converges and we obtain that $v_{\mathfrak{h}}$ is fixed under $d\pi(X)$. But this contradicts the assumption that $\mathfrak{a} \cap \mathfrak{h} = \{0\}$. □

Corollary 5.9. *The set KAH is closed in G and $(k, a, h) \mapsto kah$ is proper $K \times A \times H \to KAH$.*

6. Spaces of spherical type

Recall (see [15]) that a complex homogeneous space $G_{\mathbb{C}}/H_{\mathbb{C}}$ is said to be spherical if there exists a Borel subgroup $B_{\mathbb{C}}$ such that $B_{\mathbb{C}}H_{\mathbb{C}}$ is open in $G_{\mathbb{C}}$. The following definition is analogous. Let $Z = G/H$ be a reductive homogeneous space.

Definition 6.1. *The space Z is of* spherical type *if there exists a minimal parabolic subgroup P such that PH is open in G. If in addition $\dim(P \cap H) = 0$ then we say that Z is of* pure *spherical type.*

Note that the main intention with the concept in [15] is the classification of Gel'fand pairs. Having that intention one should add to Definition 6.1 the condition that $(M, M \cap H)$ is a Gel'fand pair. However this is not our purpose. The non-symmetric space $\mathrm{Sp}(n, 1)/\mathrm{Sp}(n)$, for example, is of spherical type but fails the Gel'fand pair condition.

It will be convenient to consider also non-minimal parabolic subgroups. If $P \subset G$ is a parabolic subgroup, we denote by $P = M_P A_P N_P$ its Langlands decomposition.

Definition 6.2. *Let $P \subset G$ be a parabolic subgroup. The pair (P, H) is called spherical if*

(1) *$M_P/(M_P \cap H)$ is compact,*
(2) *PH is open in G.*

Lemma 6.3. *If there exists a spherical pair then Z is spherical.*

Proof. Let (P, H) be a spherical pair. It follows from condition (1) that $\mathfrak{m}_P \cap \mathfrak{p} \subset \mathfrak{h}$. Hence all non-compact ideals of $\mathfrak{m}_P$ belong to $\mathfrak{h}$. Let $P^* = M^*A^*N^*$ be a minimal parabolic in M_P, then $A^*N^* \subset H$, and hence $P_0 := A_P N_P P^*$ is a minimal parabolic in G such that $P_0 H$ is open. □

6.1. Examples.

In a symmetric space Z, the minimal $\sigma\theta$-stable parabolic subgroups P satisfy (1) and (2), see [4], hence Z is of spherical type. In this case we have in addition that $P \cap H \subset M_P A_P$, and that the modular function of P is trivial on $P \cap H$.

6.1.1. *Triple spaces, Gross-Prasad spaces.* It is easily seen that the triple spaces in Corollary 5.4 and the Gross-Prasad spaces with $G_0 = \mathrm{GL}(n+1, \mathbb{R})$ are of pure spherical type. Indeed, for the chosen Cartan subspaces (see Remark 5.6) one obtains $\mathfrak{h} \oplus \mathrm{Lie}(P) = \mathfrak{g}$ with P a minimal parabolic and suitable choices of positive systems.

6.1.2. *Complex spherical spaces.* Let $G_\mathbb{C}/H_\mathbb{C}$ be a complex spherical space with open Borel orbit $B_\mathbb{C}H_\mathbb{C}$. When we regard the complex groups as real Lie groups, $G_\mathbb{C}/H_\mathbb{C}$ is of spherical type and $(B_\mathbb{C}, H_\mathbb{C})$ is a spherical pair. The complex spherical spaces have been classified (see the lists in [33] and [15]). For example, the triple space of $G_0 = \mathrm{SL}(2,\mathbb{C})$ is a complex spherical space ([33] p. 152).

6.1.3. *Real forms of spherical spaces.* Let $G_\mathbb{C}$, $H_\mathbb{C}$, $B_\mathbb{C}$ be as above, and assume that G is a quasisplit real form of $G_\mathbb{C}$. Then $B_\mathbb{C}$ is the complexification of a minimal parabolic P in G, for which PH is open. Hence G/H is of spherical type. The triple space with $G_0 = \mathrm{SL}(2,\mathbb{R})$ is obtained in this fashion.

Notice that for $n > 3$ the triple spaces with $G_0 = \mathrm{SO}_e(n,1)$ do not correspond to any spaces in 6.1.2 or 6.1.3.

6.1.4. $\mathrm{SL}(2n+1,\mathbb{C})/\mathrm{Sp}(n,\mathbb{C})$ *and* $\mathrm{SO}(2n+1,\mathbb{C})/\mathrm{GL}(n,\mathbb{C})$. According to [33] p. 143, these are complex spherical spaces. Dimension count shows they are pure. The split or quasisplit real forms are

$$\begin{aligned}
&\mathrm{SL}(2n+1,\mathbb{R})/\mathrm{Sp}(n,\mathbb{R})\\
&\mathrm{SU}(n,n+1)/\mathrm{Sp}(k,k), \quad n=2k\\
&\mathrm{SO}(n,n+1)/\mathrm{U}(k,k), \quad n=2k\\
&\mathrm{SU}(n,n+1)/\mathrm{Sp}(k,k+1), \quad n=2k+1\\
&\mathrm{SO}(n,n+1)/\mathrm{U}(k,k+1), \quad n=2k+1.
\end{aligned}$$

None of these subgroups are maximal.

6.2. **Bounds for generalized matrix coefficients.** In this paragraph we prove a fundamental bound for generalized matrix coefficients for spaces of spherical type.

To begin with we need to recall a few notions from basic representation theory. Let (π, E) be a Banach representation of G. Let us denote by p a defining norm of E. As in [10], Sect. 2, we topologize the Fréchet space of smooth vectors E^∞ by Sobolev norms Sp_s of p. These Sobolev norms are defined for any order $s \in \mathbb{R}$. We note that $Sp_{s_1} \leq Sp_{s_2}$ for $s_1 \leq s_2$. We denote by E_s the Banach-completion of (E^∞, Sp_s). If $E^{-\infty}$ denotes the continuous dual of E^∞, then we record

$$E^\infty = \bigcap_{s>0} E_s = \varprojlim E_s \quad \text{and} \quad E^{-\infty} = \bigcup_{s\in\mathbb{R}} E_s = \varinjlim E_s\,.$$

In case $s \in 2\mathbb{Z}$ is an even integer the Sobolev norms Sp_s are equivalent to norms p_s which are G-*continuous*, that is the completion of

(E^∞, p_s) defines a Banach-module for G. Hence for $k \in \mathbb{N}$ and $\eta \in E_{-2k}$ and $v \in E_{2k}$ we have

$$|\eta(\pi(g)v)| \leq p_{-2k}(\eta)p_{2k}(\pi(g)v) \tag{6.1}$$

and further

$$p_{2k}(\pi(g)v) \leq C\|g\|^l p_{2k}(v) \tag{6.2}$$

for some constants $C, l > 0$. Here $\|\cdot\|$ is a norm on G in the sense of [46], Sect. 2.

If V is a Harish-Chandra module for $(\mathfrak{g}, K)$, , then we call a Banach representation (π, E) of G a *Banach globalization* of V provided that the K-finite vectors of E are isomorphic to V as $(\mathfrak{g}, K)$-modules. The Casselman-Wallach theorem asserts that E^∞ does not depend on the particular globalization (π, E) of V and thus we may define $V^\infty := E^\infty$.

Let $Z = G/H$ be a reductive homogeneous space and V a Harish-Chandra module. For $v \in V^\infty$ and $\eta \in (V^{-\infty})^H$ an H-fixed distribution vector we denote by

$$m_{v,\eta}(gH) := \eta(\pi(g)^{-1}v), \quad (g \in G) \tag{6.3}$$

the corresponding generalized matrix coefficient. It is a smooth function on Z.

Let $\mathfrak{a} \subset \mathfrak{p}$ be maximal abelian. In this situation we fix an Iwasawa decomposition $G = KAN$ and denote by $\Pi \subset \mathfrak{a}^*$ the set of simple roots. Let $P = MAN$ denote the corresponding minimal parabolic subgroup, and $\bar{P} = \theta(P)$ its opposite.

In [46] 4.3.5 one associates to V an exponent $\Lambda \in \mathfrak{a}^*$. If $\mathfrak{g}$ is simple and π is non-trivial unitary, then it follows from the Howe-Moore-Theorem (see [43], p. 447), that

$$\Lambda \in \Big\{ \sum_{\alpha \in \Pi} c_\alpha \alpha \mid c_\alpha < 0 \Big\} . \tag{6.4}$$

Theorem 6.4. *Suppose for the minimal parabolic subgroup $P \subset G$ that $\bar{P}H$ is open in G. Let V be a Harish-Chandra module and $\eta \in (V^{-\infty})^H$. Then there exists $d \in \mathbb{N}_0$ such that for all $v \in V$ there is a constant $C_v > 0$*

$$|m_{v,\eta}(a)| \leq C_v a^\Lambda (1 + \|\log a\|)^d \qquad (a \in \overline{A^+}) . \tag{6.5}$$

Proof. The proof will be an adaption of the proof of Thm. 4.3.5 in [46]. Note that we have $m_{v,\eta}(a) = \eta(\pi(a)^{-1}v)$ whereas [46] considers $\mu(\pi(a)v)$. The linear form η will be fixed throughout the proof.

We confine ourselves to providing the key step. Our starting point is the following estimate, which follows from (6.1) and (6.2). Let (π, E)

be a Banach globalization of V, Suppose that $\eta \in E_{-l}$ for $l \in 2\mathbb{N}_0$. Then there exists $\delta \in \mathfrak{a}^*$ such that

$$|m_{v,\eta}(a)| \le a^\delta \|\eta\|_{-l}\|v\|_l \qquad (a \in \overline{A^+}) \tag{6.6}$$

where we write $\|.\|_k$ in place of p_k. If δ happens to be $\le \Lambda$ on $\mathfrak{a}^+$ we are done. Otherwise we need to improve (6.6). The key ingredient is as follows: Suppose that $v \in V$ is of the form $v = d\pi(X)u$ for some normalized positive root vector $X \in \mathfrak{g}^\alpha \subset \mathfrak{n}$ (this corresponds to the assumption $\mu \in \mathfrak{n}_F V^\sim$ in [46] p.116). As $\bar{P}H$ is open in G we can write $X = X_1 + X_2$ with $X_1 \in \mathfrak{h}$ and $X_2 \in \mathfrak{a} + \mathfrak{m} + \overline{\mathfrak{n}}$. Now observe that

$$\begin{aligned} m_{v,\eta}(a) &= \eta(\pi(a)^{-1} d\pi(X)u) = a^{-\alpha}\eta(d\pi(X)\pi(a)^{-1}u) \\ &= a^{-\alpha}\eta(d\pi(X_2)\pi(a)^{-1}u) = a^{-\alpha}\eta(\pi(a)^{-1}d\pi(\mathrm{Ad}(a)X_2)u)\,. \end{aligned}$$

As $\mathrm{Ad}(a)$ is contractive on $\mathfrak{a} + \mathfrak{m} + \overline{\mathfrak{n}}$ we can write $d\pi(\mathrm{Ad}(a)X_2)u$ as a linear combination of elements from V with a-dependent coefficients, which are bounded. We thus obtain with (6.6) an improved bound

$$|m_{v,\eta}(a)| \le C_v a^{\delta-\alpha} \qquad (a \in \overline{A^+})$$

Having seen that, the rest of the proof follows [46] p.117–118. □

Remark 6.5. (a) The theorem applies to symmetric spaces. With notation from Example 5.1.1 we let $\mathfrak{a}_q \subset \mathfrak{a}$ where $\mathfrak{a} \subset \mathfrak{p}$ is maximal abelian. Then $\mathfrak{a} = \mathfrak{a}_q \oplus \mathfrak{a}_h = (\mathfrak{a} \cap \mathfrak{q}) \oplus (\mathfrak{a} \cap \mathfrak{h})$. A proper choice of a positive chamber for $\mathfrak{a}$ yields $\mathfrak{a}_q^+ \subset \overline{\mathfrak{a}^+}$. Then PH, and hence $\bar{P}H$, is open for the corresponding minimal parabolic P. Thus (6.5) applies to all $a \in \overline{A_q^+}$. In this situation fundamental bounds as (6.5) have previously been established in [24] and [5], Thm. 14.1, and further explored in [43].
(b) We expect that there is a more quantitative version of the upper bound (6.5), namely

$$|m_{v,\eta}(a)| \le p(v) a^\Lambda (1 + \|\log a\|)^d \qquad (a \in \overline{A^+})\,. \tag{6.7}$$

for some G-continuous norm p on V.

7. Spherical type implies polar type

It is our belief that all spherical spaces are of polar type. In this section we prove this implication under certain additional hypotheses on Z.

7.1. **Algebraic type.** We will say that Z is of *algebraic type* (or *algebraic*) if there exists a reductive algebraic group $\mathbf{G}$, defined over $\mathbb{R}$, and a reductive subgroup $\mathbf{H} < \mathbf{G}$ such that $G = \mathbf{G}(\mathbb{R})$ and $H = \mathbf{H}(\mathbb{R})$. We will also write $G_\mathbb{C} = \mathbf{G}(\mathbb{C})$ and likewise we declare $K_\mathbb{C}$ and $H_\mathbb{C}$.

The crucial property that we need of an algebraic type space is the existence of a faithful finite dimensional algebraic G-module (π, V) with an H-fixed vector $v_H \in V$ such that

$$Z \to V, \quad gH \mapsto \pi(g)v_H \tag{7.1}$$

is a closed embedding (see Sect. 5.6, Th. 3 in [2]).

Assumption: From now on we assume

$$Z \text{ is of algebraic type.} \tag{7.2}$$

7.2. **Main theorem.** The main theorem in this section is

Theorem 7.1. *Let $Z = G/H$ be a reductive space of spherical type, and assume that there exists a parabolic subgroup $P \subset G$ such that*

(1) *(P, H) is spherical,*
(2) *$L := P \cap H \subset M_P$.*

Then Z is of polar type.

We shall say that the pair (P, H) is *strongly unimodular* if (1)-(2) holds. Obviously, pure spherical spaces (see Definition 6.1) are strongly unimodular. In particular it follows that the triple spaces and Gross-Prasad spaces of Example 6.1.1, as well as the spaces in 6.1.4, are all strongly spherical. Further examples are given below.

It should however be emphasized that the existence of such a pair cannot be attained for spherical spaces in general.

7.3. **Spherical unimodular type.** We prepare for the proof of Theorem 7.1 with a simple implication of Condition (2).

Definition 7.2. *Let (P, H) be a spherical pair and let $L := P \cap H$. The pair (P, H) is said to be* unimodular *if the homogeneous spaces H/L and P/L are both unimodular.*

If such pairs exist, we say that Z is of spherical unimodular type. Not all spherical pairs (P, H) are unimodular (see Example 7.4.1). The following is clear, since the modular function of P is trivial on M_P.

Lemma 7.3. *A strongly unimodular spherical pair is unimodular.*

For a symmetric space and P a minimal $\sigma\theta$-stable parabolic, the pair (P, H) is unimodular (see Example 6.1), but not necessarily strongly unimodular.

7.4. **Examples.** The first example shows that it is important to allow that P is not minimal.

7.4.1. *A case where (P,H) is not unimodular for P minimal.* Let

$$G/H = \mathrm{SL}(n+1,\mathbb{R})/\mathrm{SL}(n,\mathbb{R}).$$

This is an almost symmetric space. Let $H = \mathrm{SL}(n,\mathbb{R})$ be embedded into the right lower corner of $G = \mathrm{SL}(n+1,\mathbb{R})$, and let $P < G$ be a minimal parabolic. Then $\dim G/P = n(n+1)/2$. Specifically we choose P to be the stabilizer of the complete flag

$$\langle e_1 + e_{n+1}\rangle \subset \langle e_1 + e_{n+1}, e_2\rangle \subset \langle e_1 + e_{n+1}, e_2, e_3\rangle \subset \ldots \subset \mathbb{R}^{n+1}$$

from which we readily identify $L = P \cap H$ with the upper triangular subgroup of $\mathrm{SL}(n-1,\mathbb{R})$, where the latter group is embedded in the upper left corner of H. Thus $\dim H/L = n(n+1)/2$ and HP is open by dimension count. However H/L is not unimodular. This can be remedied by enlarging P.

If we take P to be the maximal parabolic which is the stabilizer of the line $\langle e_1+e_{n+1}\rangle$, then $L \simeq \mathrm{SL}(n-1,\mathbb{R}) \ltimes \mathbb{R}^{n-1} \subset M_P N_P$ and (P,H) is unimodular.

If instead we define P with the intermediate flag

$$\langle e_1 + e_{n+1}\rangle \subset \langle e_1 + e_{n+1}, e_2, \ldots, e_n\rangle \subset \mathbb{R}^{n+1}$$

then $L = \mathrm{SL}(n-1,\mathbb{R}) \subset M_P$ and (P,H) is strongly unimodular.

The non-symmetric space

$$G/H = \mathrm{Sp}(n+1,\mathbb{R})/\mathrm{Sp}(n,\mathbb{R}) \times \mathbb{R}$$

similarly admits a strongly unimodular pair (P,H) with a parabolic subgroup P of rank two.

7.4.2. *Gross-Prasad space with $G_0 = \mathrm{U}(p,q+1,\mathbb{F})$.* (See Section 5.1.3). A similar procedure as for Lemma 5.5 shows that this space admits a strongly unimodular pair with a minimal parabolic subgroup.

7.4.3. *Examples involving the exceptional group G_2.* In all what follows the symbol $G_2(\mathbb{C})$ denotes the simply connected complex group of type G_2. We denote by $G_2(\mathbb{R})$, resp. $U_2(\mathbb{R})$, the non-compact, resp. compact, real form of $G_2(\mathbb{C})$. All complex cases G/H below are taken from Krämer's list [33] and regarded over $\mathbb{R}$ as in Example 6.1.2.

1) $G/H = G_2(\mathbb{C})/\mathrm{SL}(3,\mathbb{C})$. In this case $\mathfrak{h}$ is maximal but non-symmetric. According to Krämer, G/H is spherical and thus there exists a Borel subgroup $P = B_{\mathbb{C}}$ of $G_2(\mathbb{C})$ such that HP is open. Then $\mathfrak{h} \cap \mathrm{Lie}(P) = \{0\}$ by dimension count and (P,H) is a pure spherical pair.

It is quite instructive to have a concrete model for Z. For that let $V_{\mathbb{R}} \simeq \mathbb{R}^7$ be the space of trace zero octonions and recall that $U_2(\mathbb{R})$ is the automorphism group of $V_{\mathbb{R}}$. If we endow $V_{\mathbb{R}}$ with an $U_2(\mathbb{R})$-invariant inner product, then $U_2(\mathbb{R})$ acts transitively on the unit sphere $S_{\mathbb{R}} \subset V_{\mathbb{R}}$ and $S_{\mathbb{R}} \simeq U_2(\mathbb{R})/\mathrm{SU}(2)$ as follows from the proof of [1], Th. 5.5. In particular, if $S_{\mathbb{C}}$ is the complex quadric and complexification of $S_{\mathbb{R}}$ in $V_{\mathbb{C}}$, then $S_{\mathbb{C}} \simeq G_2(\mathbb{C})/\mathrm{SL}(3,\mathbb{C}) = Z$. It is known that $G_2(\mathbb{R}) \cap H = \mathrm{SL}(3,\mathbb{R})$ and we obtain another pure spherical space $G_2(\mathbb{R})/\mathrm{SL}(3,\mathbb{R})$.

2) $G/H = \mathrm{SO}(7,\mathbb{C})/G_2(\mathbb{C})$. In this case $\mathfrak{h}$ is maximal, non-symmetric. Let us first recall a standard model for Z. Consider the real 8-dimensional spin representation $V_{\mathbb{R}}$ of $\mathrm{SO}(7,\mathbb{R})$ and endow $V_{\mathbb{R}}$ with an $\mathrm{SO}(7,\mathbb{R})$-invariant inner product $(,)$. We extend $(,)$ to a complex bilinear form $(,)_{\mathbb{C}}$ on $V_{\mathbb{C}}$. In $V_{\mathbb{C}}$ we choose a highest weight vector u for G and note that u is isotropic, i.e. $(u,u)_{\mathbb{C}} = 0$. The nullcone $Q_{\mathbb{C}} := \{[v] = \mathbb{C}v \in \mathbb{P}(V_{\mathbb{C}}) \mid (v,v)_{\mathbb{C}} = 0\}$ is isomorphic to G/P where P is the parabolic which fixes the line $[u] = \mathbb{C}u$. Then the Levi subgroup $M_P A_P$ of P is isomorphic to $\mathrm{GL}(3,\mathbb{C})$. Further note that $\dim_{\mathbb{C}} G = 21$ and $\dim_{\mathbb{C}} P = 15$. Now let v be a non-zero vector in $V_{\mathbb{R}}$. According to [1], Th. 5.5, the stabilizer of v in G is isomorphic to $H = G_2(\mathbb{C})$. By dimension count we have $\dim_{\mathbb{C}} L \geq 8$. We claim that there is a specific choice of v such that this intersection is isomorphic to $M = \mathrm{SL}(3,\mathbb{C})$; in particular, HP is open. For that let $u = u_1 + iu_2$ the decomposition according to $V_{\mathbb{C}} = V_{\mathbb{R}} + iV_{\mathbb{R}}$. The fact that u is isotropic means that $(u_1,u_1) = (u_2,u_2) > 0$ and $(u_1,u_2) = 0$. Take now $v = u_1$. Then L is the stabilizer of u_1 and u_2. As u_1 and u_2 are orthogonal, there is an orthogonal transformation moving u_1 to u_2. In particular L is reductive. It is a proper reductive subgroup of $G_2(\mathbb{C})$ and thus at most 8-dimensional. Our claim follows. To summarize, we constructed a strongly unimodular spherical pair (P,H) such that $P \cap H = M$. This property inherits to the real form $\mathrm{SO}(4,3)/G_2(\mathbb{R})$ of Z.

3) $G/H = \mathrm{SO}(8,\mathbb{C})/G_2(\mathbb{C})$. In this case $\mathfrak{h}$ is not symmetric and not maximal. We argue as in 2) and use a real 8-dimensional spin representation $V_{\mathbb{R}}$ of $\mathrm{SO}(8,\mathbb{R})$. The stabilizer of a highest weight vector $u \in V_{\mathbb{C}}$ is then a parabolic $P < G$ with $\dim P = 22$ and Levi subgroup $MA = \mathrm{GL}(3,\mathbb{C}) \times \mathrm{GL}(1,\mathbb{C})$ in 10 dimensions. As before we argue that $H \cap P$ is reductive and thus $L = M$. In particular, Z is strongly unimodular. Likewise the real form $\mathrm{SO}(4,4)/G_2(\mathbb{R})$ of Z is strongly unimodular as it admits a spherical pair (P,H) with $L = M$.

7.5. **Proof of Theorem 7.1.** We first study the polar decomposition on the Lie algebra level.

Lemma 7.4. *Let (P, H) be a spherical pair with Langlands decomposition $P = MAN$ of P. Then there exists $a \in A$ such that*

$$\mathfrak{k}^a + \mathfrak{a} + \mathfrak{h} = \mathfrak{g} \tag{7.3}$$

where $\mathfrak{k}^a := \mathrm{Ad}(a^{-1})(\mathfrak{k})$.

Note that in view of Sard's theorem an equivalent formulation of the conclusion is that KAH has an interior point.

Proof. Otherwise $L(a) := \mathrm{Ad}(a)\mathfrak{k} + \mathfrak{a} + \mathfrak{h}$ is a proper subspace of $\mathfrak{g}$ for all $a \in A$. For $t \mapsto a_t$ a ray tending to infinity in A^+ we note that $\lim_{t\to\infty} L(a_t) = \mathfrak{m}\cap\mathfrak{k}+\mathfrak{a}+\mathfrak{n}+\mathfrak{h}$ in the Grassmann variety of all subspaces of $\mathfrak{g}$. By (1) and (2) in Definition 6.1 we obtain $\lim_{t\to\infty} L(a_t) = \mathfrak{g}$, but as the set of subspaces of positive codimension is closed in the Grassmannian, this is a contradiction. □

Strong unimodularity of (P, H) implies the stronger decomposition

$$\mathfrak{a} \oplus (\mathfrak{k}^a + \mathfrak{h}) = \mathfrak{g}. \tag{7.4}$$

This results from the following lemma, where we replace $L \subset M$ by the weaker assumption $L \subset MA$.

Lemma 7.5. *Assume $L \subset MA$ and let $a \in A$ satisfy (7.3). Then*

$$\mathfrak{a} \cap (\mathfrak{k}^a + \mathfrak{h}) = \mathfrak{a} \cap \mathfrak{h} \quad \textit{and} \quad \mathfrak{k}^a \cap \mathfrak{h} = \mathfrak{m} \cap \mathfrak{k} \cap \mathfrak{h}.$$

Proof. Since (P, H) is a spherical pair we have

$$\mathfrak{h} + (\mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}) = \mathfrak{g} \tag{7.5}$$

As $\mathfrak{m}\cap\mathfrak{p} \subset \mathfrak{h}$ we have $\mathfrak{l} = (\mathfrak{m}\cap\mathfrak{p}) \oplus (\mathfrak{m}\cap\mathfrak{k}\cap\mathfrak{h}) \oplus (\mathfrak{a}\cap\mathfrak{h})$. Comparing the excess dimensions in (7.3) and (7.5) and using that $\dim \mathfrak{k} = \dim(\mathfrak{m} \cap \mathfrak{k}) + \dim \mathfrak{n}$ we infer

$$\dim[\mathfrak{a} \cap (\mathfrak{k}^a + \mathfrak{h})] + \dim(\mathfrak{k}^a \cap \mathfrak{h}) = \dim(\mathfrak{m} \cap \mathfrak{k} \cap \mathfrak{h}) + \dim(\mathfrak{a} \cap \mathfrak{h}).$$

Since a centralizes $\mathfrak{m}$ we have $\mathfrak{m} \cap \mathfrak{k} \cap \mathfrak{h} \subset \mathfrak{k}^a \cap \mathfrak{h}$, and we are done. □

Lemma 7.6. *Assume that (P, H) is a strongly unimodular spherical pair. Then $G = KAH$.*

Proof. Given in Appendix B. □

Theorem 7.1 now follows from Corollary 5.9.

7.6. **Strongly spherical.** To facilitate the exposition in the next section we need the following concept, which combines polar and spherical types compatibly.

Definition 7.7. *A reductive homogeneous space $Z = G/H$ is of* strong spherical type *if the following holds. There exists $\mathfrak{a} \subset \mathfrak{p}$ maximal abelian and minimal parabolic subgroups $P_1, \dots, P_l \supset A$ such that*

(1) *Z is of polar type and $G = KAH$;*
(2) *Z is of spherical type with $\bar{P}_j H$ is open for all $1 \leq j \leq l$;*
(3) *$\bigcup_{j=1}^{l} \overline{A^+(P_j)}(A \cap H) = A$.*

Symmetric spaces are strongly spherical, see Remark 6.5. We believe that

$$\text{spherical type} \;\Rightarrow\; \text{strong spherical type.} \tag{7.6}$$

but again we shall only prove this under the additional assumption of strong unimodularity.

Example 7.8. Consider G/H with

(1) $G = \mathrm{GL}(n+1, \mathbb{R})$ and $H = \mathrm{GL}(n, \mathbb{R})$ for $n \geq 0$.
(2) $G = \mathrm{U}(p, q+1, \mathbb{C})$ and $H = \mathrm{U}(p, q, \mathbb{C})$ for $p + q \geq 2$.

These spaces are (rank one) symmetric modulo the center of G. It follows easily that they are strongly spherical.

Corollary 7.9. *If Z is of strongly unimodular spherical type, then it is strongly spherical.*

Proof. Let (P, H) be a strongly unimodular spherical pair. We have seen in Lemma 7.6 that $G = KA_PH$.

Let P^* be a minimal parabolic in M_P and define a minimal parabolic P_0 in G by $P_0 = A_PN_PP^*$ as in the proof of Lemma 6.3. Then $P \supset P_0$ and (P_0, H) is a spherical pair. Let $\mathfrak{a}_0 \subset \mathfrak{p}$ be the corresponding Cartan subspace.

Note that the sphericality of (P_0, H) is an open condition for P_0. This implies that (kP_0k^{-1}, H) is spherical for generic $k \in K$. By choosing k sufficiently generic we can thus ensure that $(kwP_0w^{-1}k^{-1}, H)$ is spherical for all Weyl group elements w of $\mathfrak{a}_0$. We now replace P by its k-conjugate kPk^{-1}, while noting that in view of Lemma B.2, this conjugate is again unimodular. Likewise we replace P_0 and A_0 by their k-conjugates, and obtain thus that (Q, H) is spherical for all minimal parabolics Q with $A_Q = A_0$,

Let $Q_1, \dots, Q_l$ be an enumeration of the parabolic subgroups Q for which $A_Q = A_P$. Each Q_j contains one or more minimal parabolics with split part $\mathfrak{a}_0$, of which we choose one and denote it by $\bar{P}_j$. Then $\bar{P}_j H$ is open for each j. The result follows. □

8. Property (I)

Let $Z = G/H$ be of reductive type. We introduce an integrability condition for matrix coefficients on Z. It has some similarity with Kazdan's Property (T).

We denote by $\widehat{G}$ the unitary dual of G and by $\widehat{G}_s$ the subset of K-spherical representations. With $\mathfrak{a} \subset \mathfrak{p}$ a maximal abelian subspace, and W the corresponding Weyl group, we can identify $\widehat{G}_s$ with a subset of $\mathfrak{a}_{\mathbb{C}}^*/W$.

Definition 8.1. *We say that a reductive homogeneous space $Z = G/H$ has* Property (I) *if for all $\pi \in \widehat{G}_s$ and $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ the stabilizer H_η of η is reductive in G and there exists $1 \leq p < \infty$ such that*

$$m_{v,\eta} \in L^p(G/H_\eta), \tag{8.1}$$

for all $v \in \mathcal{H}_\pi^\infty$.

The following lemma shows that it suffices to have (8.1) for $v \in \mathcal{H}_\pi^K$.

Lemma 8.2. *Let $(\pi, \mathcal{H}_\pi)$ be irreducible unitary, and let $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ and $1 \leq p < \infty$. The following statements are equivalent:*

1. *$m_{v,\eta} \in L^p(Z)$ for all $v \in \mathcal{H}_\pi^\infty$.*
2. *$m_{v,\eta} \in L^p(Z)$ for all K-finite vectors in $\mathcal{H}_\pi$.*
3. *$m_{v,\eta} \in L^p(Z)$ for some K-finite vector $v \neq 0$.*

Proof. Let V be the Harish-Chandra module of $(\pi, \mathcal{H}_\pi)$, i.e. the space of K-finite vectors. According to Harish-Chandra, V is an irreducible $(\mathfrak{g}, K)$-module. The map $v \mapsto m_{v,\eta}$ is equivariant $V \to C^\infty(Z)$.

We first establish "(3) ⇒ (2)". Let $v \in V$ be non-zero with $m_{v,\eta} \in L^p(Z)$, then v generates V, and (2) is equivalent with the statement that $m_{v,\eta} \in L^p(Z)^\infty$.

Let E be the closed G-invariant subspace of $L^p(Z)$ generated by $m_{v,\eta}$. As the left action on $L^p(Z)$ is a Banach representation, the same holds for E. The Casimir element $\mathcal{C}$ acts by a scalar on V, hence it acts (in the distribution sense) on E by the same scalar. It follows that all K-finite vectors in E are smooth for the Laplacian Δ associated to $\mathcal{C}$ (see (11.2) below). Thus any K-finite vector of E belongs to $E^\infty \subset L^p(Z)^\infty$ by [10], Prop. 2.13.

Finally "(2) ⇒ (1)" follows from the Casselman-Wallach globalization theorem (see [10]), which implies that the map $v \mapsto m_{v,\eta}$, $V \to L^p(Z)$, extends to $\mathcal{H}_\pi^\infty \to L^p(Z)^\infty$. □

In the definition of property (I) we have to take into account that the stabilizer H_η inflates H. To discuss this efficiently it is useful to have appropriate notion of factorization for Z.

8.1. **Factorization.** Given a reductive homogeneous space $Z = G/H$. If there exists a reductive subgroup $H^* \supset G$ such that $\mathfrak{h} \subsetneq \mathfrak{h}^* \subsetneq \mathfrak{g}$, then we call Z *factorizable*, set $Z^* = G/H^*$ and call

$$Z \mapsto Z^*, \quad gH \mapsto gH^*$$

a *factorization* of Z.

Example 8.3. 1) Irreducible symmetric spaces are not factorizable. In fact, if $(\mathfrak{g}, \mathfrak{h})$ is a symmetric pair and $\mathfrak{h} \subsetneq \mathfrak{h}^* \subsetneq \mathfrak{g}$ a factorization, then $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$, $\mathfrak{h} = \mathfrak{h}_1 \times \mathfrak{h}_2$ and $\mathfrak{h}^* = \mathfrak{h}_1 \times \mathfrak{g}_2$ for some decomposition of $\mathfrak{g}$.
2) Suppose that $H < G$ is a proper reductive subgroup. Then $Z = G \times H/\operatorname{diag}(H)$ is factorizable with $H^* = H \times H$.
3) Let $G^n := G \times \cdots \times G$ denote the direct product of n copies of G. Suppose that G is simple. The homogeneous space $Z = G^n/\operatorname{diag}(G)$ with the diagonal subgroup $\operatorname{diag}(G) = \{(g, \dots, g) | g \in G\}$ is factorizable if and only if $n > 2$. For example, for $n = 3$, we can take $H^* = \{(g_1, g_2, g_2) \mid g_1, g_2 \in G\} \simeq G \times G$ and permutations thereof.
4) Let $G/H = \mathrm{SO}(8, \mathbb{C})/G_2(\mathbb{C})$ as in Example 7.4.3.3. The symmetric space $G/H^* = \mathrm{SO}(8, \mathbb{C})/\mathrm{SO}(7, \mathbb{C})$ is a factorization.

Having the notion of factorization in the back of our mind we resume our discussion of property (I) in the context of these examples.

Example 8.4. 1) Suppose that $Z = G \times H/\operatorname{diag}(H)$ with G simple. Irreducible spherical unitary representations of $G \times H$ are of the form $\pi = \pi_1 \otimes \pi_2$ where $\pi_1 \in \widehat{G}_s$ and $\pi_2 \in \widehat{H}_s$. Moreover π is H-spherical if and only if π_2^* is contained in $\pi_1|_H$. If π_2 was the trivial representation, then the stabilizer of η inflates to $H_\eta = H \times H$.
2) Suppose that $Z = G \times G \times G/\operatorname{diag}(H)$ with G simple. Irreducible spherical representations of $G \times G \times G$ are tensor products $\pi = \pi_1 \otimes \pi_2 \otimes \pi_3$. If π is non-trivial and H-spherical and if one constituent, say π_1, is trivial, then $H_\eta = \{(g_1, g_2, g_2) \mid g_1, g_2 \in G\}$.

8.2. **Main Theorem on Property** (I)**.** We can now state one of the main results of the paper. Recall that a space with the properties in Definition 7.7 is called strongly spherical.

Theorem 8.5. *Let $Z = G/H$ be reductive and suppose that Z is strongly spherical. Then Z has Property* (I)*.*

Proof. Note that any factorization Z^* of Z will be strongly spherical, once Z is strongly spherical. Proceeding by induction, we may thus assume that all factorizations Z^* have property (I). We may also assume that $\mathfrak{h}$ contains no non-trivial ideal of $\mathfrak{g}$.

Let $\pi \in \widehat{G}_s$ be non-trivial and let $f = m_{v,\eta}$ with $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ and v a K-fixed vector. In view of Lemma 8.2 it is sufficient to show that $f \in L^p(G/H_\eta)$ for some $1 \leq p < \infty$.

Since $G = KAH$ the invariant measure μ_Z on Z can be expressed in $K \times A$-coordinates as

$$\int_Z \phi(z) d\mu_Z(z) = \int_K \int_{A/(A\cap H)} \phi(ka \cdot z_0) J(a)\; da\; dk \qquad (\phi \in C_c(Z))\,.$$

The Jacobian is readily verified to be the absolute value of a polynomial in a^α, $\alpha \in \Sigma$. If the exponent Λ of π satisfies (6.4), it follows from Theorem 6.4 that $|f|^p$ becomes integrable over $A^+(P_j)/(A \cap H)$, for sufficiently large p. Hence $f \in L^p(Z)$. If the exponent Λ does not satisfy (6.4), then G is not simple and π is 1 on some non-trivial normal subgroup in G. This subgroup is contained in H_η but not in H, hence $Z \to G/H_\eta$ is a factorization. Now the assertion follows from our inductive hypothesis. □

Corollary 8.6. *Let Z be a homogeneous space of reductive type. Assume that Z is symmetric or one of the following spherical spaces*

1. *triple spaces $G^3/\operatorname{diag}(G)$ with $G = \mathrm{SO}_e(1,n)$ and $n \geq 2$,*
2. *Gross-Prasad spaces $G \times H/\operatorname{diag}(H)$, see Section 5.1.3,*
3. *the spaces discussed in Examples 6.1.4, 7.4.1, 7.4.3, and 7.8.*

Then Z has Property (I).

We emphasize that this is by no means an exhaustive list of spaces which can be treated by the present methods.

9. Counting lattice points, I: Preliminaries

In this section we let G/H be a homogeneous space of reductive type. We assume that there exists a lattice (a discrete subgroup with finite covolume) $\Gamma \subset G$ such that $\Gamma_H := \Gamma \cap H$ is a lattice in H. We normalize Haar measures on G and H such that:

- $\mathrm{vol}(G/\Gamma) = 1$.
- $\mathrm{vol}(H/\Gamma_H) = 1$.

To simplify the presentation we assume that G is linear semi-simple. Let us emphasize that this is just a matter of convenience because it allows us to skip integration over a possibly non-compact center.

Our concern is with the double fibration

$$Z := G/H \leftarrow G/\Gamma_H \rightarrow Y := G/\Gamma\,.$$

Fibrewise integration yields transfer maps from functions on Z to functions on Y and vice versa. In more precision,

$$L^\infty(Y) \to L^\infty(Z),\ \phi \mapsto \phi^H;\ \phi^H(gH) := \int_{H/\Gamma_H} \phi(gh\Gamma)\ d(h\Gamma_H) \tag{9.1}$$

and we record that this map is contractive, i.e

$$\|\phi^H\|_\infty \leq \|\phi\|_\infty \qquad (\phi \in L^\infty(Y))\,. \tag{9.2}$$

Likewise we have

$$L^1(Z) \to L^1(Y),\ f \mapsto f^\Gamma;\ f^\Gamma(g\Gamma) := \sum_{\gamma\in\Gamma/\Gamma_H} f(g\gamma H)\,, \tag{9.3}$$

which is contractive, i.e

$$\|f^\Gamma\|_1 \leq \|f\|_1 \qquad (f \in L^1(Z))\,. \tag{9.4}$$

Unfolding with respect to the double fibration yields, in view of our normalization of measures, the following adjointness relation:

$$\langle f^\Gamma, \phi\rangle_{L^2(Y)} = \langle f, \phi^H\rangle_{L^2(Z)} \tag{9.5}$$

for all $\phi \in L^\infty(Y)$ and $f \in L^1(Z)$. Let us note that (9.5) applied to $|f|$ and $\phi = \mathbf{1}_Y$ readily yields (9.4).

9.1. **Distance function.** We recall that on G one defines a K-bi-invariant distance function $g \mapsto \|g\|_G$ by $\|k \exp X\|_G = \|X\|$ for $k \in K$ and $X \in \mathfrak{p}$. In geometric terms, $\|g\|_G$ is the Riemannian distance of $Kg \in K\backslash G$ from the origin. It is well known that

$$\|xy\|_G \leq \|x\|_G + \|y\|_G$$

for all $x, y \in G$. The norm $\|g\|$ used in (6.2) relates as $\log \|g\| = c\|g\|_G$ with $c > 0$.

On Z we recall the polar decomposition (4.1) and define similarly $\|[k, X]\|_Z = \|X\|$ for $k \in K$ and $X \in \mathfrak{q} \cap \mathfrak{p}$. We record the following

Lemma 9.1. *Let $z = gH \in Z$. Then $\|z\|_Z = \inf_{h\in H} \|gh\|_G$.*

Proof. (cf [31], Lemma 5.4). It suffices to prove that $\|\exp(X)h\|_G \geq \|X\|$ for $X \in \mathfrak{q} \cap \mathfrak{p}$, $h \in H$, and by Cartan decomposition of H, we may assume $h = \exp(T)$ with $T \in \mathfrak{h} \cap \mathfrak{p}$. Thus we have reduced to the statement that

$$\|\exp(X)\exp(T)\|_G \geq \|\exp(X)\|_G$$

for $X \perp T$ in $\mathfrak{p}$. This follows from the fact that the sectional curvatures of $K\backslash G$ are ≤ 0 (see [28], p. 73). □

In particular, it follows that

$$\|gz\|_Z \leq \|z\|_Z + \|g\|_G \quad (z \in Z, g \in G). \tag{9.6}$$

9.2. **Generalized balls.** The problem of counting lattice points in Z leads to a natural question of exhibiting families of compact subsets that exhaust Z. We introduce here two families of generalized balls.

We define the *intrinsic ball* of radius $R > 0$ on Z by

$$B_R := \{z \in Z \mid \|z\|_Z < R\}.$$

Note that the notation B_R was used differently in Section 4. It follows from (9.6) that for each $r > 0$ there exists a $K \times K$-invariant neighborhood U of $\mathbf{1}$ in G such that

$$gB_R \subset B_{R+r} \tag{9.7}$$

for all $g \in U$ and all $R > 0$.

We are interested in the volume of B_R, and shall write it $|B_R|$. It follows easily from the lower bounds in Section 4 that $|B_R| \nearrow \infty$ as $R \to \infty$ (in fact $|B_R| \geq CR^{\dim \mathfrak{p}\cap\mathfrak{q}}$). On the other hand, we have the following upper bound.

Lemma 9.2. *There exists a constant $c > 0$ such that*

$$|B_{R+r}| \leq e^{cr}|B_R| \tag{9.8}$$

for all $R \geq 1, r \geq 0$.

Proof. It follows from the formula for the invariant integral with respect to (4.1) that

$$|B_R| = \int_{X\in\mathfrak{q}\cap\mathfrak{p}, \|X\|<R} J(X)\, dX$$

where $J(X) = J_{\mathbf{1}}(X)$.

Hence it suffices to prove that there exists $c > 0$ such that

$$\int_0^{R+r} J(tX)t^{l-1}\, dt \leq e^{cr} \int_0^R J(tX)t^{l-1}\, dt$$

for all $X \in \mathfrak{q} \cap \mathfrak{p}$ with $\|X\| = 1$. Here $l = \dim \mathfrak{q} \cap \mathfrak{p}$. Equivalently, the function

$$R \mapsto e^{-cR} \int_0^R J(tX)t^{l-1}\, dt$$

is decreasing, or by differentiation,

$$J(RX)R^{l-1} \leq c \int_0^R J(tX)t^{l-1}\, dt$$

for all R. The latter inequality is established in [23, Lemma A.3] with c independent of X. □

We write $\mathbf{1}_R \in L^1(Z)$ for the characteristic function of B_R and deduce from the definitions and (9.5):

- $\mathbf{1}_R^\Gamma(e\Gamma) = N_R(\Gamma, Z) := \#\{\gamma \in \Gamma/\Gamma_H \mid \gamma.z_0 \in B_R\}$.
- $\|\mathbf{1}_R^\Gamma\|_{L^1(G/\Gamma)} = |B_R|$.

9.2.1. *Other balls.* A second family of generalized balls can be constructed by embedding $\iota : Z \hookrightarrow V$ in a vector space V by choosing a finite dimensional representation (ρ, V) of G with H the stabilizer of a vector in V (see Lemma 5.7). We fix a K-invariant norm ρ on V and define

$$B_{R,\rho} = \{z \in Z : \rho(\iota(z)) < R\}$$

to be the intersection of the ball of radius R in V with our subvariety Z. We write $|B_{R,\rho}|$ for the volumes of these balls in Z, and note that properties of these functions of R are established in [23, Appendix 1].

Formulas similar to the bulleted hold for the corresponding characteristic functions $\mathbf{1}_{R,\rho}$.

9.3. **Factorization of balls.** Let us call $Z = G/H$ *rigid* if any factorization is compact modulo the center $\mathcal{Z}$ of G, i.e. if $Z \to Z^*$ is a factorization of Z, then $H^*\mathcal{Z}/H\mathcal{Z}$ is compact. Note that all irreducible symmetric spaces are rigid.

In case Z is not rigid, then, with regard to the lattice count, we limit our consideration to balls which satisfy natural additional properties. In the sequel U_R either stands for B_R or $B_{R,\rho}$.

Let Z be non-rigid and $F \to Z \to Z^*$ be a factorization with fiber space $F := H^*/H$. We write U_R^* for the image of U_R under the factorization map, i.e. $U_R^* = U_R H^*/H^*$. Further we set $U_R^F := F \cap U_R$.

For a compactly supported bounded measurable function ϕ on Z we define the fiberwise integral

$$\phi^F(gH^*) := \int_{H^*/H} \phi(gh^*)\ d(h^*H)$$

and recall the integration formula

$$\int_Z \phi(gH)\ d(gH) = \int_{Z^*} \phi^F(gH^*)\ d(gH^*) \tag{9.9}$$

under appropriate normalization of measures. Consider the characteristic function $\mathbf{1}_R$ of U_R and note that its fiber average $\mathbf{1}_R^F$ is supported in the compact ball U_R^*. We say that the family of balls $(U_R)_{R>0}$ *factorizes well* provided for all compact subsets $Q \subset G$

$$\lim_{R\to\infty} \frac{\sup_{g\in Q} 1_R^F(gH^*)}{|U_R|} = 0 \tag{9.10}$$

holds for all factorizations. For our balls $U_R = B_R, B_{R,\rho}$ one readily obtains for all compact subsets Q an $R_0 = R_0(Q) > 0$ such that

$$\sup_{g\in Q} 1_R^F(gH^*) \leq |U_{R+R_0}^F| \,.$$

by (9.7). Thus the balls U_R factorize well provided

$$\lim_{R\to\infty} \frac{|U_{R+R_0}^F|}{|U_R|} = 0\,.$$

for all $R_0 > 0$. In practice, this appears always to be satisfied.

Example 9.3. (a) Let us consider geometric balls in a triple space $Z = G \times G \times G/\Delta(G)$. Write B_R^G for the geometric ball in G. Then

$$B_R = (B_R^G \times B_R^G \times B_R^G)\Delta(G) \subset Z$$

by Lemma 9.1. We identify Z with $G \times G$ via the coordinate map

$$G \times G \to Z, \quad (g_1, g_2) \mapsto (g_1, g_2, \mathbf{1})\Delta(G)$$

and within this identification we have $B_R = (B_R^G \times B_R^G)\Delta(B_R^G) \subset G \times G$. Consider $H^* = \{(g, h, h) \mid g, h \in G\}$ and note that H^*/H identifies with $G \times \{\mathbf{1}\} \subset Z$. In particular $B_R^F = B_{2R}^G$. Thus $(B_R)_{R>0}$ factorizes well provided

$$\lim_{R\to\infty} \frac{\mathrm{vol}_G\, B_{2R+R_0}^G}{\mathrm{vol}_{G\times G}(B_R^G \times B_R^G)\Delta(B_R^G)} = 0$$

for all $R_0 > 0$. With Fubini the denominator is $\mathrm{vol}_G\, B_{3R}^G$ and thus the balls factorize well.

(b) More generally, all geometric balls for the spaces listed in the introduction factorize well. These cases fit in a more general pattern. Suppose the factorization $H^*/H \to G/H \to G/H^*$ has the following properties:

1. (1) There exist a strongly unimodular spherical pair (P, H) for G such that $G = KAH$.
2. (2) With $A^* = A \cap H^*$, $M^* = M \cap H^*$, $N^* = N \cap H^*$ one gets a spherical pair (P^*, H) for H^* with $P^* = M^*A^*N^*$ such that $H^* = K^*A^*H$.

Under this conditions the geometric balls B_R factorize well. Indeed let us denote by $J : \mathfrak{a} \to \mathbb{R}_{\geq 0}$ the Jacobian of the polar map $K \times A \to Z$ and likewise we define J^*. Then $|B_R| = \int_{\{\|X\|\leq R\}} J(X)\, dX$ and an according formula for the volume of the fibered balls $B_R^F = B_R \cap H^*$. Now for $X \in \mathfrak{a}^+$ away from the walls one readily gets the asymptotic behavior $J(X) \sim e^{2\rho_{\mathfrak n}(X)}$ and likewise $J^*(X) \sim e^{2\rho_{\mathfrak n^*}(X)}$. The assertion follows.

9.4. **Weak asymptotics.** Under the assumption that $Z = G/H$ is symmetric and H/Γ_H is compact the main result in [20] reads as

$$(\mathrm{MT}_\rho) \qquad N_{R,\rho}(\Gamma, Z) \sim |B_{R,\rho}| \quad (R \to \infty).$$

The compactness assumption was later dropped in [22] where an ergodic proof was presented.

For spaces with property (I) and Y compact we show in the following section that (MT_ρ), as well as the corresponding result for the intrinsic balls B_R,

$$(\mathrm{MT}) \qquad N_R(\Gamma, Z) \sim |B_R| \quad (R \to \infty),$$

can be deduced from VAI, Theorem 2.2.

With notation from (9.3) we set

$$F_R^\Gamma := \frac{1}{|B_R|}\mathbf{1}_R^\Gamma, \qquad \text{and} \quad F_{R,\rho}^\Gamma := \frac{1}{|B_{R,\rho}|}\mathbf{1}_{R,\rho}^\Gamma \qquad (R > 0)\,.$$

We shall concentrate in verifying the following limits of weak type:

$$(\mathrm{wMT}_\rho) \qquad \langle F_{R,\rho}^\Gamma, \phi\rangle \to \int_Y \bar\phi\, d\mu_Y \quad (R \to \infty), \qquad (\forall \phi \in C_0(Y))$$

$$(\mathrm{wMT}) \qquad \langle F_R^\Gamma, \phi\rangle \to \int_Y \bar\phi\, d\mu_Y \quad (R \to \infty), \qquad (\forall \phi \in C_0(Y))\,.$$

Lemma 9.4. $(\mathrm{wMT}_\rho) \Rightarrow (\mathrm{MT}_\rho)$ *and* $(\mathrm{wMT}) \Rightarrow (\mathrm{MT})$.

Proof. The first implication is shown in [20] Lemma 2.3 for Z a symmetric space. The proof is based on a property of the action of G on the family of balls under consideration, derived in an appendix of [20]. The same property is derived for general reductive homogenous spaces in [23, Appendix 1] (in [22] the relevant property is denominated as *well-roundedness*). It follows that (wMT_ρ) implies (MT_ρ) in general. The second implication is proved similarly from (9.7) and (9.8). □

10. Counting lattice points, II: Main term analysis

In this section we will establish main term counting under the mandate of property (I) and Y being compact.

In Section 9.3 we defined the notions of non-rigid spaces and well factorizable balls therein. To have a uniform notation, we understand that both families $(B_R)_{R>0}$ and $(B_{R,\rho})_{R>0}$ are well factorizable in case Z is rigid.

10.1. **Main Theorem on main term counting.**

Theorem 10.1. *Let G be reductive and H a closed subgroup such that $Z = G/H$ is of reductive type. Suppose that Y is compact and Z admits* (I). *Then the following assertion hold true:*

(1) *If $(B_R)_{R>0}$ factorizes well, then* (wMT) *and* (MT) *hold.*
(2) *If $(B_{R,\rho})_{R>0}$ factorizes well, then* (wMT$_\rho$) *and* (MT$_\rho$) *hold.*

By combining with Theorem 8.5 one obtains main term counting for symmetric spaces, as in [20], but also for the geometric balls for the spaces listed in the introduction.

The proof is based on the following lemma. For a function space $\mathcal{F}(Y)$ consisting of integrable functions on Y we denote by $\mathcal{F}(Y)_o$ the subspace of functions with vanishing integral.

Proposition 10.2. *Let $Z = G/H$ be of reductive type. Assume that there exists a dense subspace $\mathcal{A}(Y) \subset C_b(Y)_o^K$ such that*

$$\phi^H \in C_0(Z) \quad \textit{for all } \phi \in \mathcal{A}(Y)\,. \tag{10.1}$$

Then (wMT) *and* (wMT$_\rho$) *hold.*

Proof. We concentrate on (wMT) and establish it for $\phi \in C_b(Y)$. As $C_b(Y) = C_b(Y)_o \oplus \mathbb{C}\mathbf{1}_Y$, and (wMT) is trivial for ϕ a constant, it suffices to establish

$$\langle F_R^\Gamma, \phi\rangle \to 0 \qquad (\phi \in C_b(Y)_o)\,. \tag{10.2}$$

We will show (10.2) is valid for $\phi \in \mathcal{A}(Y)$. By density, as F_R^Γ is K-invariant and has norm one in $L^1(Y)$, this will finish the proof.

Let $\phi \in \mathcal{A}(Y)$ and let $\epsilon > 0$. By the unfolding identity (9.5) we have

$$\langle F_R^\Gamma, \phi\rangle_{L^2(Y)} = \frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H\rangle_{L^2(Z)}. \tag{10.3}$$

Using (10.1) we choose $K_\epsilon \subset Z$ compact such that $|\phi^H(z)| < \epsilon$ outside of K_ϵ. Then

$$\frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H\rangle_{L^2(Z)} = \int_{K_\epsilon} + \int_{Z-K_\epsilon} \frac{\mathbf{1}_R(z)}{|B_R|}\phi^H(z)\, d\mu_Z(z)\,.$$

By (9.2), the first term is bounded by $\frac{|K_\epsilon| \|\phi\|_\infty}{|B_R|}$, which is $\leq \epsilon$ for R sufficiently large. As the second term is bounded by ϵ for all R, we obtain (10.2). Hence (wMT) holds. Exactly the same argument gives (wMT$_\rho$). $\square$

Remark 10.3. It is possible to replace (10.1) by a weaker requirement: Suppose that $\mathcal{A}(Y) = \sum_{j\in J} \mathcal{A}(Y)_j$ where $\mathcal{A}(Y)_j$ is a subspace for which the $\phi^H \in \mathcal{A}(Y)_j$ factorize to an invariant function $\phi^{H_j^*}$. Suppose that

$$\phi^{H_j^*} \in C_0(Z_j^*) \qquad (\phi \in \mathcal{A}(Y)_j) \tag{10.4}$$

holds for all $j \in J$. Then the conclusion in Lemma 10.2 is still valid, provided the relevant balls B_R or $B_{R,\rho}$ factorize well. In fact, using (9.9) the last part of the proof modifies to (with $H^* = H_j^*$ to simplify notation):

$$\begin{aligned}\frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H\rangle_{L^2(Z)} &= \frac{1}{|B_R|}\langle \mathbf{1}_R^F, \phi^{H^*}\rangle_{L^2(Z^*)} = \\ &= \int_{K_\epsilon^*} + \int_{Z^*-K_\epsilon^*} \frac{\mathbf{1}_R^F(z)}{|B_R|}\phi^{H^*}(z)\, d\mu_Z(z)\,.\end{aligned}$$

As $\|1_R^F\|_{L^1(Z^*)} = |B_R|$, the second term is bounded by ϵ for all R, and the first term we get as small as we wish with (9.10)

10.2. **The space $\mathcal{A}(Y)$.** We now construct a specific subspace $\mathcal{A}(Y) \subset C_b(Y)_0^K$ for and verify condition (10.1).

As Y is compact, the abstract Plancherel-theorem reads:

$$L^2(G/\Gamma)^K \simeq \bigoplus_{\pi\in\widehat{G}_s} \mathcal{V}_{\pi,\Gamma}$$

where $\mathcal{V}_{\pi,\Gamma} \subset (\mathcal{H}_\pi^{-\infty})^\Gamma$ is a subspace of finite dimension accounting for multiplicities. If we denote the Fourier transform by $f \mapsto f^\wedge$ then the corresponding inversion formula is given by

$$f = \sum_\pi m_{v_\pi, f^\wedge(\pi)} \tag{10.5}$$

with $v_\pi \in \mathcal{H}_\pi$ normalized K-fixed and $f^\wedge(\pi) \in \mathcal{V}_{\pi,\Gamma}$. The matrix coefficients for Y are defined as in (6.3), and the sum in (10.5) is required to include multiplicities.

Note that $L^2(Y) = L^2(Y)_o \oplus \mathbb{C}\cdot \mathbf{1}_Y$. We define $\mathcal{A}(Y) \subset L^2(Y)_o^K$ to be the dense subspace of functions with finite Fourier support. Then $\mathcal{A}(Y) \subset L^2(Y)^\infty$ is dense and since $C^\infty(Y)$ and $L^2(Y)^\infty$ are topologically isomorphic, it follows that $\mathcal{A}(Y)$ is dense in $C(Y)_o^K$ as required.

Lemma 10.4. *Assume that Y is compact and Z has* (I)*, and define $\mathcal{A}(Y)$ as above. Then condition* (10.1) *holds if Z is rigid, and otherwise* (10.4).

Proof. Assume first that Z is rigid. It is no loss of generality to assume that the center of G is compact. The map $\phi \mapsto \phi^H$ corresponds on the spectral side to a map $(\mathcal{H}_\pi^{-\infty})^\Gamma \to (\mathcal{H}_\pi^{-\infty})^H$, which can be constructed as follows.

As H/Γ_H is compact, we can define

$$\Lambda_\pi : (\mathcal{H}_\pi^{-\infty})^\Gamma \to (\mathcal{H}_\pi^{-\infty})^H, \quad \Lambda_\pi(\eta) = \int_{H/\Gamma_H} \eta \circ \pi(h^{-1}) \, d(h\Gamma_H) \tag{10.6}$$

by $\mathcal{H}_\pi^{-\infty}$-valued integration: the defining integral is understood as integration over a compact fundamental domain $F \subset H$ with respect to the Haar measure on H; as the integrand is continuous and $\mathcal{H}_\pi^{-\infty}$ is a complete locally convex space, the integral converges in $\mathcal{H}_\pi^{-\infty}$.

Let $\phi \in \mathcal{A}(Y)$, then it follows from (10.5) that

$$\phi^H = \sum_{\pi \neq \mathbf{1}} m_{v_\pi, \Lambda_\pi(\phi^\wedge(\pi))} \, . \tag{10.7}$$

As Z has property (I), the finite sum in (10.7) belongs to $L^p(G/H)$ for some $1 \leq p < \infty$ (to be precise modulo the center). We conclude that $\phi^H \in L^p(G/H)$, and from Lemma 8.2 that $\phi^H \in L^p(G/H)^\infty$. In view of VAI (Theorem 2.2), the claim now follows.

In case Z is not rigid we let J denote the set of all factorizations $Z^* \to Z$ (including also $Z^* = Z$) and define $\mathcal{A}(Y)_j$ for $j \in J$ accordingly to accommodate summands in (10.7) for which $H_\eta = H_j^*$. □

10.2.1. *The example of Eskin-McMullen.* We end this section by treating another class of non-symmetric spaces $Z = G/H$ for which the lattice density along the intrinsic balls satisfies (MT). The theory includes the spaces which were introduced by Eskin and McMullen [22] as counterexamples to (MT$_\rho$), see Example 10.6 below.

We are interested in spaces $Z = G/H$ for which the normalizer $N_K(H)$ of H in K is not contained in H. In this case

$$Z' := G/H', \qquad H' := HN_K(H) \, ,$$

is a factorization of Z.

Proposition 10.5. *Let $Z' = G/H'$ be as above. If Y is compact, Z' admits* (I)*, and the balls B_R in Z factorize well, then* (wMT) *and* (MT) *hold for $Z = G/H$.*

Proof. Letting $N_K(H)$ act on $Z = G/H$ from the right we obtain an action of the extended group $\tilde{G} = G \times N_K(H)$ on Z such that

$$Z \simeq \tilde{G}/\tilde{H}, \quad \tilde{H} = \{(hk, k) \mid h \in H, k \in N_K(H)\}.$$

The $\tilde{G}$-invariant measure on Z agrees with the G-invariant, by uniqueness of the latter. Furthermore, as $\mathfrak{q} \cap \mathfrak{p}$ and its norm are preserved in the adjoint action of $N_K(H)$, the intrinsic balls B_R of Z for $\tilde{G}$ and G

are identical (note that in contrast the other type of balls $B_{R,\rho}$ are not defined for $\tilde{G}$, unless the H-fixed vector of ρ is also fixed by $N_K(H)$).

We claim that $Z' = G/H'$ admits (I) if and only if $Z = \tilde{G}/\tilde{H}$ admits it. Let $\tilde{K} = K \times N_K(H)$, then an irreducible unitary representation $\pi \otimes \delta$ of $\tilde{G}$ has a non-trivial $\tilde{K}$-fixed vector only if δ is trivial. This reduces the verification of (I) for $\tilde{G}/\tilde{H}$ to representations π of G with $(\mathcal{H}_\pi^{-\infty})^{H'} \neq 0$, and the claim follows. The proposition now follows from Theorem 10.1 ☐

We end this section by providing some examples where Proposition 10.5 can be applied.

Example 10.6. Consider the case of $G = \mathrm{SL}(2,\mathbb{C})$ and $H < G$ the subgroup of real diagonal matrices. Hence $H' = HN_K(H)$ is the subgroup of all diagonal matrices, so that G/H' is a symmetric space

The lattice Γ is a conjugate of $\mathrm{SL}(2,\mathbb{Z}[i])$ so that H/Γ_H is compact. In view of Proposition 10.5 and the result of the next section, it follows that (MT) holds.

This example was considered in [22] and [23], where it is shown that (MT_ρ) fails for a particular finite dimensional representation ρ. This shows that (wMT_ρ) is false in general for the situation in Theorem 10.5.

Example 10.7. Let G be a real reductive group and let $H = \{e\}$. Then $Z = G$ is a homogeneous space for the left action of G. In this case $N_K(H) = K$ and $Z' = G/K$ is symmetric, hence admits (I). Hence (MT) holds for every cocompact lattice in G.

11. Counting lattice points, III: Error term analysis

In this section we assume that Z is symmetric. However, we wish to point out that this assumption enters only in the key Lemma 11.5. For simplicity we also assume that G is simple. In addition we request that the cycle $H/\Gamma_H \subset Y$ is compact. This technical condition ensures that the integration map

$$\Lambda_\pi : (\mathcal{H}_\pi^{-\infty})^\Gamma \to (\mathcal{H}_\pi^{-\infty})^H, \quad \Lambda_\pi(\eta) = \int_{H/\Gamma_H} \eta \circ \pi(h^{-1})\, d(h\Gamma_H) \tag{11.1}$$

considered in (10.6) is defined. In the sequel we use the Plancherel theorem (see [27])

$$L^2(G/\Gamma)^K \simeq \int_{\widehat{G}_s}^{\oplus} \mathcal{V}_{\pi,\Gamma}\, d\mu(\pi)\,,$$

where $\mathcal{V}_{\pi,\Gamma} \subset (\mathcal{H}_\pi^{-\infty})^\Gamma$ is a finite dimensional subspace and of constant dimension on each connected component in the continuous spectrum (parametrization by Eisenstein series), and where the Plancherel measure μ has support

$$\widehat{G}_\mu := \operatorname{supp}(\mu) \subset \widehat{G}_s.$$

The first error term for (wMT) can be expressed by

$$\operatorname{err}(R, \Gamma) := \sup_{\substack{\phi \in C_b(Y) \\ \|\phi\|_\infty \leq 1}} |\langle F_R^\Gamma - \mathbf{1}_Y, \phi\rangle| \qquad (R > 0),$$

and our goal is to give an upper bound for $\operatorname{err}(R, \Gamma)$ as a function of R.

Remark 11.1. In the literature results are sometimes stated with respect to the pointwise error term $\operatorname{err}_{pt}(R, \Gamma) = |F_R^\Gamma(\mathbf{1}) - |B_R||$. Note that

$$\operatorname{err}_{pt}(R, \Gamma) \leq \sup_{\substack{\phi \in L^1(Y) \\ \|\phi\|_1 \leq 1}} |\langle F_R^\Gamma - \mathbf{1}_Y, \phi\rangle| \qquad (R > 0).$$

The Sobolev estimate $\|\phi\|_\infty \leq C\|\phi\|_{1,k}$, for K-invariant functions ϕ on Y and with $k = \dim Y/K$ the Sobolev shift, then relates these error terms.

According to the decomposition $C_b(Y) = C_b(Y)_o \oplus \mathbb{C}\mathbf{1}_Y$ we decompose functions as $\phi = \phi_o + \phi_1$ and obtain

$$\operatorname{err}(R, \Gamma) = \sup_{\substack{\phi \in C_b(Y) \\ \|\phi\|_\infty \leq 1}} |\langle F_R^\Gamma, \phi_o\rangle| = \sup_{\substack{\phi \in C_b(Y) \\ \|\phi\|_\infty \leq 1}} \frac{|\langle \mathbf{1}_R, \phi_o^H\rangle|}{|B_R|}.$$

Further, from $\|\phi_o\|_\infty \leq 2\|\phi\|_\infty$ we obtain that $\operatorname{err}(R, \Gamma) \leq 2\operatorname{err}_1(R, \Gamma)$ with

$$\operatorname{err}_1(R, \Gamma) := \sup_{\substack{\phi \in C_b(Y)_o \\ \|\phi\|_\infty \leq 1}} \frac{|\langle \mathbf{1}_R, \phi^H\rangle|}{|B_R|}.$$

11.1. **L^p bounds on generalized matrix coefficients of H-distinguished representations.** Let $(\pi, \mathcal{H}_\pi)$ be a non-trivial unitary irreducible spherical representation of G. Then by Lemma 8.5 it follows that there exists a smallest index $1 \leq p_H(\pi) < \infty$ such that all K-finite generalized matrix coefficients $m_{v,\eta}$ with $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ (see (6.3)) belong to $L^p(Z)$ for any $p > p_H(\pi)$.

The representation π is said to be H-distinguished if $(\mathcal{H}_\pi^{-\infty})^H \neq \{0\}$. Note that if π is not H-distinguished then $p_H(\pi) = 1$. We say that π is H-tempered if $p_H(\pi) = 2$.

Given a lattice $\Gamma \subset G$ we define

$$p_H(\Gamma) := \sup\{p_H(\pi) : \pi \in \widehat{G}_\mu, \pi \neq 1\}$$

and record the following.

Lemma 11.2. $p_H(\Gamma) < \infty$.

Proof. This follows from [18] in the cases where G has property (T). The remaining cases are $\mathrm{SO}_e(n,1)$ and $\mathrm{SU}(n,1)$ (up to covering), for which the result is well known. □

Remark 11.3. (a) It is not difficult to relate the finiteness of $p_H(\Gamma)$ to the notion of "spectral gap" which is common in ergodic theory (see [21] or [30]).
(b) In many cases one expects a Ramanujan-Selberg property, see [44]. In particular one expects $p_H(\Gamma) = 2$ in these cases.

Recall the Cartan-Killing form κ on $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ and choose a basis $X_1, \ldots, X_l$ of $\mathfrak{k}$ and $X_1', \ldots, X_s'$ of $\mathfrak{p}$ such that $\kappa(X_i, X_j) = -\delta_{ij}$ and $\kappa(X_i', X_j') = \delta_{ij}$. With that data we form the standard Casimir element

$$\mathcal{C} := -\sum_{j=1}^{l} X_j^2 + \sum_{j=1}^{s} (X_j')^2 \in \mathcal{U}(\mathfrak{g}) .$$

Set $\Delta_K := \sum_{j=1}^{l} X_j^2 \in \mathcal{U}(\mathfrak{k})$ and obtain the commonly used Laplace element

$$\Delta = \mathcal{C} + 2\Delta_K \in \mathcal{U}(\mathfrak{g}) \tag{11.2}$$

which acts on $Y = G/\Gamma$ from the left.

Let $d \in \mathbb{N}$. For $1 \le p \le \infty$, it follows from [10], Section 2, that Sobolev norms on $L^p(Y)^\infty \subset C^\infty(Y)$ can be defined by

$$||f||_{p,2d}^2 = \sum_{j=0}^{d} ||\Delta^j f||_p^2 .$$

Basic spectral theory allows to declare $\| \cdot \|_{p,d}$ for any $d \ge 0$.

Let us define

$$s := \dim \mathfrak{p} = \dim G/K = \dim \Gamma \backslash G/K$$

and

$$r := \dim \mathfrak{a} = \mathrm{rank}_{\mathbb{R}}(G/K) ,$$

where $\mathfrak{a} \subset \mathfrak{p}$ is maximal abelian.

Proposition 11.4. *Let $Z = G/H$ and $p > p_H(\Gamma)$. The map*

$$\mathrm{Av_H} : C_b^\infty(Y)_o^K \to L^p(Z)^K ; \mathrm{Av_H}(\phi) = \phi^H$$

is continuous. More precisely, for all

(1) $k > s + 1$ *if* Y *is compact.*

(2) $k > \frac{r+1}{2}s + 1$ *if* Y *is non-compact and* Γ *is arithmetic*

there exists a constant $C = C(p,k) > 0$ *such that*

$$\|\phi^H\|_p \leq C\|\phi\|_{\infty,k} \qquad (\phi \in C_b^\infty(Y)_o^K)$$

Before we prove the proposition we need a fiberwise estimate. Given $\pi \in \widehat{G}_\mu$ we consider the functions $\phi = \phi_\pi \in C_b^\infty(Y)^K$ which are generalized matrix coefficients

$$\phi_\pi(g\Gamma) := m_{v,\eta}(g\Gamma) = \eta(\pi(g^{-1})v), \quad (g \in G) \tag{11.3}$$

with $v \in \mathcal{H}_\pi^\infty$ and $\eta \in (\mathcal{H}_\pi^{-\infty})^\Gamma$.

Let $1 \leq p < \infty$. Let us say that a subset $\Lambda \subset \widehat{G}_s$ is $L^p(Z)$*-bounded* provided that $m_{v,\eta} \in L^p(Z)$ for all $\pi \in \Lambda$ and $v \in \mathcal{H}_\pi^\infty$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$. The proof of the following Lemma is postponed to Appendix C.

Lemma 11.5. *Suppose that* $\Lambda \subset \widehat{G}_s$ *is* $L^p(Z)$*-bounded for some* $1 \leq p < \infty$*. Then there exists* $C > 0$ *such that*

$$\|m_{v,\eta}\|_p \leq C\|m_{v,\eta}\|_\infty$$

for all $\pi \in \Lambda$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ *and* $v \in \mathcal{H}_\pi^K$.

As a consequence we obtain that:

Lemma 11.6. *Let* $p > p_H(\Gamma)$*. Then there exists* $C > 0$ *such that*

$$\|\phi_\pi^H\|_p \leq C\|\phi_\pi\|_\infty$$

for all $\pi \in \widehat{G}_\mu$, $v \in \mathcal{H}_\pi^K$, $\eta \in (\mathcal{H}_\pi^{-\infty})^\Gamma$*, with* ϕ_π *given by* (11.3)*.*

Proof. Recall from (9.2), that integration is a bounded operator from $L^\infty(Y) \to L^\infty(Z)$. Hence the assertion follows from the previous lemma. $\square$

Proof. (Proposition 11.4) For all $\pi \in \widehat{G}$ the operator $d\pi(\mathcal{C})$ acts as a scalar which we denote by

$$|\pi| := d\pi(\mathcal{C})\,.$$

Let $\phi \in C_b^\infty(Y)_o^K$ and write $\phi = \phi_d + \phi_c$ for its decomposition in discrete and continuous Plancherel parts. We assume first that $\phi = \phi_d$.

In case Y is compact we have Weyl's law: There is a constant $c_Y > 0$ such that

$$\sum_{|\pi| \leq R} m(\pi) \sim c_Y R^{s/2} \qquad (R \to \infty)\,.$$

Here $m(\pi) = \dim \mathcal{V}_{\pi,\Gamma}$. We conclude that

$$\sum_\pi m(\pi)(1 + |\pi|)^{-k} < \infty \tag{11.4}$$

for all $k > s/2+1$. In case Y is non-compact, we let $\widehat{G}_{\mu,d}$ be the the discrete support of the Plancherel measure. Then assuming Γ is arithmetic, the upper bound in [29] reads:

$$\sum_{\substack{\pi\in\widehat{G}_{\mu,d}\\ |\pi|\leq R}} m(\pi)\leq c_Y R^{rs/2} \qquad (R>0)\,.$$

For $k > rs/2+1$ we obtain (11.4) as before.

As ϕ is in the discrete spectrum we decompose it as $\phi=\sum_\pi \phi_\pi$ and obtain with Lemma 11.6

$$\|\phi^H\|_p \leq \sum_\pi \|\phi_\pi^H\|_p \leq C_p \sum_\pi \|\phi_\pi\|_\infty\,.$$

In these sums representations occur according to their multiplicities $m(\pi)$. The last sum we estimate as follows:

$$\begin{aligned}\sum_\pi \|\phi_\pi\|_\infty &= \sum_\pi (1+|\pi|)^{-k/2}(1+|\pi|)^{k/2}\|\phi_\pi\|_\infty\\ &\leq C\sum_\pi (1+|\pi|)^{-k/2}\|\phi_\pi\|_{\infty,k}\end{aligned}$$

with $C>0$ a constant depending only on k. Applying the Cauchy-Schwartz inequality combined with (11.4) we obtain

$$\|\phi^H\|_p \leq C\Big(\sum_\pi \|\phi_\pi\|_{\infty,k}^2\Big)^{\frac12}$$

with $C>0$ (we allow universal positive constants to change from line to line).

To finish the proof we apply the Sobolev lemma on $K\backslash G$. Here Sobolev norms are defined by the central operator $\mathcal{C}$, whose action agrees with the left action of Δ. It follows that $\|f\|_\infty \leq C\|f\|_{2,k_1}$ with $k_1>\frac{s}{2}$ for K-invariant functions f on G. This gives

$$\|\phi^H\|_p\leq C(\sum_\pi ||\phi_\pi||_{2,k+k_1}^2)^{\frac12} = C||\phi||_{2,k+k_1}\leq C||\phi||_{\infty,k+k_1}$$

which proves the proposition for the discrete spectrum.

If $\phi=\phi_c$ belongs to the continuous spectrum, where multiplicities are bounded, the proof is simpler. Let μ_c be the restriction of the Plancherel measure to the continuous spectrum. As this is just Euclidean measure on r-dimensional space we have

$$\int_{\widehat{G}_s}(1+|\pi|)^{-k}\, d\mu_c(\pi)<\infty \tag{11.5}$$

if $k > r/2$. We assume for simplicity in what follows that $m(\pi) = 1$ for all $\pi \in \operatorname{supp}\mu_c$. As $\sup_{\pi\in\operatorname{supp}\mu_c} m(\pi) < \infty$ the proof is easily adapted to the general case.

Let

$$\phi = \int_{\widehat{G}_s} \phi_\pi \; d\mu_c(\pi).$$

As $\|\phi^H\|_\infty \le \|\phi\|_\infty$ we conclude with Lemma 11.6, (11.5) and Fubini's theorem that

$$\phi^H = \int_{\widehat{G}_s} \phi_\pi^H \; d\mu_c(\pi)$$

and, by the similar chain of inequalities as in the discrete case

$$\|\phi^H\|_p \le C\|\phi\|_{\infty,k+k_1}$$

with $k > \frac{r}{2}$ and $k_1 > \frac{s}{2}$. This concludes the proof. □

11.2. **Smooth versus non-smooth counting.** Like in the classical Gauss circle problem one obtains much better estimates for the remainder term if one uses a smooth cutoff. Let $\alpha \in C_c^\infty(G)$ be a non-negative test function with normalized integral. Set $\mathbf{1}_{R,\alpha} := \alpha * \mathbf{1}_R$ and define

$$\operatorname{err}_\alpha(R,\Gamma) := \sup_{\substack{\phi\in C_b(Y)_o^K\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_{R,\alpha}, \phi^H\rangle|}{|B_R|}\,.$$

Lemma 11.7. *Let $k > s+1$ if Y is compact and $k > \frac{r+1}{2}s+1$ otherwise. Let $p > p_H(\Gamma)$. Then there exists $C > 0$ such that*

$$\operatorname{err}_\alpha(R,\Gamma) \le C\|\alpha\|_{1,k}|B_R|^{-\frac{1}{p}}$$

for all $R \ge 1$ and all $\alpha \in C_c^\infty(G)$.

Proof. First note that

$$\langle \mathbf{1}_{R,\alpha}, \phi^H\rangle = \langle \mathbf{1}_{R,\alpha}, (-\mathbf{1}+\Delta)^{k/2}(-\mathbf{1}+\Delta)^{-k/2}\phi^H\rangle\,.$$

With $\psi = (-\mathbf{1}+\Delta)^{-k/2}\phi$ we have $\|\psi\|_{\infty,k} \le C\|\phi\|_\infty$ for some $C > 0$. We thus obtain

$$\operatorname{err}_\alpha(R,\Gamma) \le C \sup_{\substack{\psi\in C_b(Y)_o^K\\ \|\psi\|_{\infty,k}\le 1}} \frac{|\langle \mathbf{1}_{R,\alpha}, (-\mathbf{1}+\Delta)^{k/2}\psi^H\rangle|}{|B_R|}\,.$$

Moving $(-\mathbf{1}+\Delta)^{k/2}$ to the other side we get with Proposition 11.4 and Hölder's inequality that

$$\begin{aligned}\operatorname{err}_\alpha(R,\Gamma) &\leq C \sup_{\substack{\psi\in C_b(Y)_o^K\\ \|\psi\|_{\infty,k}\leq 1}} \frac{|\langle(-\mathbf{1}+\Delta)^{k/2}\alpha * \mathbf{1}_R, \psi^H\rangle|}{|B_R|}\\ &\leq C\frac{\|(-\mathbf{1}+\Delta)^{k/2}\alpha * \mathbf{1}_R\|_q}{|B_R|}\end{aligned}$$

where q is the conjugate exponent satisfying $\frac{1}{p}+\frac{1}{q}=1$. Finally,

$$\|(-\mathbf{1}+\Delta)^{k/2}\alpha * \mathbf{1}_R\|_q \leq C\|\alpha\|_{1,k}\|\mathbf{1}_R\|_q$$

and the Lemma follows. □

Finally we have to compare $\operatorname{err}_1(R,\Gamma)$ with $\operatorname{err}_\alpha(R,\Gamma)$. For that we note that

$$|\operatorname{err}_1(R,\Gamma) - \operatorname{err}_\alpha(R,\Gamma)| \leq \sup_{\substack{\phi\in C_b(Y)_o^K\\ \|\phi\|_\infty\leq 1}} \frac{|\langle \mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma, \phi\rangle|}{|B_R|}\,.$$

Suppose that $\operatorname{supp}\alpha \subset B_\epsilon^G$ for some $\epsilon > 0$ with the superscript G indicating that we take the balls in the reductive space "G". Then (9.7) implies that $\mathbf{1}_{R,\alpha}$ is supported in $B_{R+\epsilon}$, and hence

$$\begin{aligned}|\langle \mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma, \phi\rangle| &\leq \|\mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma\|_1\\ &\leq \|\mathbf{1}_{R,\alpha} - \mathbf{1}_R\|_1\\ &\leq |B_{R+\epsilon}|^{\frac{1}{2}}\|\mathbf{1}_{R,\alpha} - \mathbf{1}_R\|_2\\ &\leq |B_{R+\epsilon}|^{\frac{1}{2}}|B_{R+\epsilon}\backslash B_R|^{\frac{1}{2}}\,.\end{aligned}$$

From (9.8) we have

$$|B_{R+\epsilon}\backslash B_R| \leq C\epsilon|B_R| \qquad (R\geq 1, \epsilon<1)\,.$$

Thus we obtain that

$$|\operatorname{err}_1(R,\Gamma) - \operatorname{err}_\alpha(R,\Gamma)| \leq C\epsilon^{\frac{1}{2}}\,.$$

Combining this with the estimate in Lemma 11.7 we arrive at the existence of $C>0$ such that

$$\operatorname{err}_1(R,\Gamma) \leq C(\epsilon^{-k}|B_R|^{-\frac{1}{p}} + \epsilon^{\frac{1}{2}})$$

for all $R\geq 1$ and all $0<\epsilon<1$. The minimum of the function $\epsilon\mapsto\epsilon^{-k}c+\epsilon^{1/2}$ is attained at $\epsilon=(2kc)^{\frac{2}{2k+1}}$ and thus we get:

Theorem 11.8. *The first error term* $\mathrm{err}(R,\Gamma)$ *for the lattice counting problem on* $Z = G/H$ *can be estimated as follows: for all* $p > p_H(\Gamma)$ *and* $k > s+1$ *for* Y *compact, resp.* $k > \frac{r+1}{2}s+1$ *otherwise, there exists a constant* $C = C(p,k) > 0$ *such that*

$$\mathrm{err}(R,\Gamma) \leq C|B_R|^{-\frac{1}{(2k+1)p}}$$

for all $R \geq 1$.

Remark 11.9. The point where we loose essential information is in the estimate (11.4) where we used Weyl's law. In the moment pointwise multiplicity bounds are available the estimate would improve. To compare the results with Selberg on the hyperbolic disc, let us assume that $p_H(\Gamma) = 2$. Then with $r = 1$ and $s = 2$ our bound is $\mathrm{err}(R,\Gamma) \leq C_\epsilon |B_R|^{-\frac{1}{14}+\epsilon}$ while Selberg showed $\mathrm{err}(R,\Gamma) \leq C_\epsilon |B_R|^{-\frac{1}{3}+\epsilon}$.

12. Counting lattice points IV: Triple spaces

In this final section we reveal an explicit study of the triple case $Z = \mathrm{PSl}(2,\mathbb{R})^3/\operatorname{diag}(\mathrm{PSl}(2,\mathbb{R}))$. We present an alternative approach towards (I) and establish the key Lemma 11.5, which implies error term counting.

Let us define subgroups of $G_0 = \mathrm{PSl}(2,\mathbb{R}) \simeq \mathrm{SO}_e(2,1)$:

$$\begin{aligned} A_0 &= \left\{ a_t = \begin{pmatrix} t & 0 \\ 0 & \frac{1}{t} \end{pmatrix} \mid t > 0 \right\}, \\ B_0 &= \left\{ b_s = \begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix} \mid s \in \mathbb{R} \right\}, \\ N_0 &= \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R} \right\}, \end{aligned}$$

and $P_0 = A_0 N_0$. We note that G/H is known to be a multiplicity one space, that is, the space $(\mathcal{H}^{-\infty})^H$ of H-fixed distribution vectors is at most one-dimensional for all $\pi \in \widehat{G}$ (see [41]). This property is closely related to the fact that there is only one open $P = P_0^3$-orbit on Z. Indeed, if

$$w_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad s_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

and $z_0 = (\mathbf{1}, w_0, s_0)$, then

$$P z_0 H$$

is open and dense in G: to see that, observe that $P\backslash G = \mathbb{P}^1(\mathbb{R})^3$ and that H acts transitively on all transversal triples.

This suggests the following definition

$$A = A_0 \times A_0 \times B_0$$

and, indeed, we have already established in Proposition 5.3 that $G = KAH$.

12.1. **The invariant measure on Z.** For a general semisimple group G_0 we note that the map

$$G_0 \times G_0 \to Z, \quad (g_1, g_2) \mapsto (g_1, g_2, \mathbf{1})H \tag{12.1}$$

is a $G_0 \times G_0$-equivariant diffeomorphism. Accordingly the invariant measure on Z identifies with the Haar measure on $G_0 \times G_0$.

For $G_0 = \mathrm{PSl}(2, \mathbb{R})$ we can actually compute the Haar-measure in terms of the KAH-coordinates. Let us identify A with $\mathbb{R}^3$ via the map

$$(t_1, t_2, s) \mapsto (a_{t_1}, a_{t_2}, b_s)\,.$$

Then

$$d\mu_Z(gH) = J(t_1, t_2, s)\, dk\, dt_1\, dt_2\, ds\,. \tag{12.2}$$

Lemma 12.1. $J(t_1, t_2, s) = |\sinh(2(t_1 - t_2)) \cosh(2s)|$

Proof. On $G_0 \times G_0$ we use the formula for integration in HAK coordinates for the symmetric space $\mathrm{diag}(G_0)\backslash G_0 \times G_0 = G_0$: The map

$$G_0 \times A_0 \times K_0 \times K_0 \to G_0 \times G_0$$

defined by

$$(g, a_t, k_1, k_2) \mapsto (g, g)(a_{t/2}, a_{-t/2})(k_1, k_2)$$

is a parametrization and the Haar measure on $G_0 \times G_0$ writes as

$$|\sinh(2t)| dg\ da_t\ dk_1\ dk_2\,.$$

Note that the "half" is swallowed in the identification $(x, y) \mapsto x^{-1}y$ of $\mathrm{diag}(G_0)\backslash G_0 \times G_0$ with G_0.

Further we decompose $G_0 = \mathrm{diag}(G_0)$ by means of the KAH coordinates for the symmetric space G_0/A_0: the map

$$K_0 \times B_0 \times A_0 \to G_0, \quad (k, b_s, a) \mapsto kb_s a$$

gives

$$dg = \cosh(2s)\ dk\ db_s\ da\,.$$

Combining, we have the coordinates $(kb_s a_{u+t/2}k_1, kb_s a_{u-t/2}k_2)$ on $G_0 \times G_0$, with Jacobian $|\sinh(2(t_1 - t_2)) \cosh(2s)|$ where we have written $t_1 = u + t/2$ and $t_2 = u - t/2$.

Finally we identify $G_0 \times G_0$ with $\mathrm{diag}(G_0)\backslash G_0 \times G_0 \times G_0$ by

$$(g_1, g_2) \mapsto \mathrm{diag}(G_0).(g_1, g_2, \mathbf{1})$$

so that the above coordinates correspond to

$$\operatorname{diag}(G_0).(a_{t_1}, a_{t_2}, b_{-s})(k_1, k_2, k^{-1})$$

□

12.2. **The key lemma.** As H is not maximal we first need a modified notion for an L^p-bounded subset $\Lambda \subset \widehat{G}_s\backslash\{\mathbf{1}\}$ (compare the definition given above Lemma 11.5): we need to request that $m_{v,\eta} \in L^p(G/H_\eta)$ instead of $m_{v,\eta} \in L^p(G/H)$. Having said that, the key lemma reads:

Lemma 12.2. *Let $Z = G_0^3/\operatorname{diag}(G_0)$ and $G_0 = \mathrm{PSl}(2,\mathbb{R})$. Suppose that $\Lambda \subset \widehat{G}_s$ is L^p-bounded for some $1 \leq p < \infty$. Then there exists $C > 0$ such that*

$$\|m_{v,\eta}\|_p \leq C\|m_{v,\eta}\|_\infty$$

for all $\pi \in \Lambda$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ and $v \in \mathcal{H}_\pi^K$.

Proof. We write $\Lambda_0 := \Lambda \cap [\widehat{G_{0s}}\backslash\{\mathbf{1}\}]^3$ and $\Lambda_1 := \Lambda\backslash\Lambda_0$.

Let $(\pi, \mathcal{H}_\pi)$ be a non-trivial K-spherical unitary representation of G. Then $\pi = \pi_1 \otimes \pi_2 \otimes \pi_3$ with each factor a K_0-spherical unitary representation of G_0. We assume that π has non-trivial H-fixed distribution vectors, then at least two of the factors π_i are non-trivial.

Let v_i be normalized K_0-fixed vectors of π_i and set $v = v_1 \otimes v_2 \otimes v_3$. Since Z is a multiplicity one space, the functional $I \in (\mathcal{H}_\pi^{-\infty})^H$ is unique up to scalars. Our concern is with the L^p-integrability of the generalized matrix coefficient $f_\pi := m_{v,I}$:

$$f_\pi(g_1, g_2, g_3) := I(\pi_1(g_1)^{-1}v_1 \otimes \pi_2(g_2)^{-1}v_2 \otimes \pi_3(g_3)^{-1}v_3)\,,$$

when π belongs to Λ. Consider first the case where $\pi \in \Lambda_1$, i.e. one π_i is trivial, say π_1. Then $\pi_2 = \pi_3^*$. Viewed as a function on $G \times G \simeq Z$ by means of (12.1), f_π becomes

$$f_\pi(g_1, g_2) = \langle \pi_2(g_2)v_2, v_2\rangle\,.$$

This function is constant on the first factor and by assumption L^p-integrable on the second. As the key lemma is satisfied for the symmetric space G_0/K_0, the assertion follows for $\pi \in \Lambda_1$.

Suppose now that $\pi \in \Lambda_0$, i.e. all π_i are non-trivial. In order to analyze f_π we use $G = KAH$ and thus assume that $g = a = (a_1, a_2, b) \in A$. We work in the compact model of $\mathcal{H}_{\pi_i} = L^2(\mathbb{S}^1)$ and use the explicit model for I in [11]: for f_1, f_2, f_3 smooth functions on the circle one has

$$I(f_1 \otimes f_2 \otimes f_3) = \frac{1}{(2\pi)^3}\int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi} f_1(\theta_1)f_2(\theta_2)f_3(\theta_3)\cdot \\ \cdot K(\theta_1, \theta_2, \theta_3)\; d\theta_1 d\theta_2 d\theta_3\,,$$

where

$$K(\theta_1,\theta_2,\theta_3) = |\sin(\theta_2-\theta_3)|^{(\alpha-1)/2}|\sin(\theta_1-\theta_3)|^{(\beta-1)/2}|\sin(\theta_1-\theta_2)|^{(\gamma-1)/2}.$$

In this formula one has $\alpha = \lambda_1 - \lambda_2 - \lambda_3$, $\beta = -\lambda_1 + \lambda_2 - \lambda_3$ and $\gamma = -\lambda_1 - \lambda_2 + \lambda_3$ where $\lambda_i \in i\mathbb{R} \cup (-1/2, 1/2)$ are the standard representation parameters of π_i. We note that when the λ_i are real, we have a free choice of their sign, so that it always can be arranged that K is integrable.

Returning to our analysis of f_π we now take $f_1(\theta_1) = [\pi_1(a_{t_1})v_1](\theta_1)$, $f_2(\theta_2) = [\pi_2(a_{t_2})v_2](\theta_2)$ and $f_3 = [\pi_3(b_s)v_3](\theta_3)$. Then

$$f_1(\theta_1) = \frac{1}{(t_1^2 + \sin^2\theta_1(\frac{1}{t_1^2} - t_1^2))^{\frac{1}{2}(1+\lambda_1)}}$$

and likewise formulas for f_2 and f_3.

A simple computation then yields the existence of constants $c_i = c_i(\pi) > 0$, $C = C(\pi) > 0$ depending on π only through the distance of $\operatorname{Re}\lambda_i$ to the trivial representation, such that

$$|f_\pi(a_{t_1}, a_{t_2}, b_s)| \leq C\frac{1}{[\cosh\log t_1]^{c_1}\cdot[\cosh\log t_2]^{c_2}\cdot[\cosh s]^{c_3}}.$$

This bound is essentially sharp. Hence, in view of the explicit form (12.2) of the invariant measure in KAH-coordinates, it follows that

$$\inf_{\pi\in\Lambda_0} c_i(\pi) > 0 \qquad \text{and} \qquad \inf_{\pi\in\Lambda_0} C(\pi) > 0.$$

In particular we get that

$$\sup_{\pi\in\Lambda_0} \|f_\pi\|_p < \infty.$$

On the other hand for $g = \mathbf{1} = (\mathbf{1},\mathbf{1},\mathbf{1})$, the value $f_\pi(\mathbf{1})$ is obtained by applying I to the constant function $\mathbf{1} = \mathbf{1}\otimes\mathbf{1}\otimes\mathbf{1}$. This value has been computed explicitly by Bernstein and Reznikov in [12]. The result is that $I(\mathbf{1})$ is quotient of Γ-functions which, after Stirling approximation, yields a lower bound for $f_\pi(\mathbf{1})$:

$$\inf_{\pi\in\Lambda_0} |f_\pi(\mathbf{1})| > 0.$$

As $\|f_\pi\|_\infty \geq |f_\pi(\mathbf{1})|$ the proof is finished. □

Remark 12.3. Let us emphasize that the same proof is applicable to the case where $G_0 = \mathrm{SO}_e(1,n)$ (see [17] for the generalization of [12]).

With the key Lemma 12.2 we obtain as in Section 11 (see Theorem 11.8) the following error term bound:

Theorem 12.4. *Let $Z = G_0^3/\operatorname{diag}(G_0)$ for $G_0 = \mathrm{SO}_e(1,n)$ and assume that H/Γ_H is compact. Then the first error term $\operatorname{err}(R,\Gamma)$ for the lattice counting problem on $Z = G/H$ can be estimated as follows: for all $p > p_H(\Gamma)$ there exists a $C = C(p) > 0$ such that*

$$\operatorname{err}(R,\Gamma) \leq C|B_R|^{-\frac{1}{(6n+3)p}}$$

for all $R \geq 1$.

Appendix A. Proof of Theorem 2.2

In this appendix we prove that VAI does not hold on any homogeneous space $Z = G/H$ of G, which is not of reductive type. We maintain the assumption that G is a real reductive Lie group and establish the following result.

Proposition A.1. *Assume that $H \subset G$ is a closed connected subgroup such that $Z = G/H$ is unimodular and not of reductive type. Then for all $1 \leq p < \infty$ there exists an unbounded function $f \in L^p(Z)^\infty$. In particular,* VAI *does not hold.*

The idea is to show that there is a compact ball $B \subset G$ and a sequence $(g_n)_{n\in\mathbb{N}}$ such that

- $Bg_nz_0 \cap Bg_mz_0 = \emptyset$ for $n \neq m$.
- $\operatorname{vol}_Z(Bg_nz_0) \leq e^{-n}$ for all $n \in \mathbb{N}$.

Out of these data it is straightforward to construct a smooth L^p-function which does not vanish at infinity.

Before we give a general proof we first discuss the case of unipotent subgroups. The argument in the general case, although more technical, will be modeled after that.

A.1. **Unipotent subgroups.** Let $H = N$ be a unipotent subgroup, that is, $\mathfrak{n} := \mathfrak{h}$ is an ad-nilpotent subalgebra of $[\mathfrak{g},\mathfrak{g}]$. Now, the situation where $\mathfrak{n}$ is normalized by a particular semi-simple element is fairly straightforward and we shall begin with a discussion of that case.

If $X \in \mathfrak{g}$ is a real semi-simple element, i.e., $\operatorname{ad} X$ is semi-simple with real spectrum, then we denote by $\mathfrak{g}_X^\lambda \subset \mathfrak{g}$ its eigenspace for the eigenvalue $\lambda \in \mathbb{R}$, and by $\mathfrak{g}_X^\pm$ the sum of these eigenspaces for λ positive/negative. We record the triangular decomposition

$$\mathfrak{g} = \mathfrak{g}_X^+ + \mathfrak{z}_\mathfrak{g}(X) + \mathfrak{g}_X^- .$$

Here $\mathfrak{z}_\mathfrak{g}(X) =: \mathfrak{g}_X^0$ is the centralizer of X in $\mathfrak{g}$.

Lemma A.2. *Assume that $\mathfrak{n}$ is normalized by a non-zero real semisimple element $X \in \mathfrak{g}$ such that $\mathfrak{n} \subset \mathfrak{g}_X^+$. Set $a_t := \exp(tX)$ for all $t \in \mathbb{R}$.*

Let $B \subset G$ be a compact ball around $\mathbf{1}$. Then there exists $c > 0$ and $\gamma > 0$ such that

$$\mathrm{vol}_Z(Ba_t z_0) = c \cdot e^{t\gamma} \qquad (t \in \mathbb{R})$$

Proof. Let $A = \exp \mathbb{R}X$ and note that A normalizes N. Thus for all $a \in A$ the prescription

$$\mu_{Z,a}(Bz_0) := \mu_Z(Baz_0) \qquad (B \subset G \text{ measurable})$$

defines a G-invariant measure on Z. By the uniqueness of the Haar measure we obtain that

$$\mu_{Z,a} = J(a)\mu_Z$$

where $J : A \to \mathbb{R}_0^+$ is the group homomorphism $J(a) = \det \mathrm{Ad}(a)|_{\mathfrak{n}}$. All assertions follow. □

Having obtained this volume bound we can proceed as follows. Let us denote by χ_k the characteristic function of $Ba_{-k}z_0 \subset Z$. We claim that the non-negative function

$$\chi := \sum_{k\in\mathbb{N}} k\chi_k \tag{A.1}$$

lies in $L^p(G/H)$. In fact

$$\|\chi\|_p \leq \sum_{k\in\mathbb{N}} k\|\chi_k\|_p \leq c\sum_{k\in\mathbb{N}} ke^{-\gamma k/p}\,.$$

Finally we have to smoothen χ: For that let $\phi \in C_c(G)^\infty$ with $\phi \geq 0$, $\int_G \phi = 1$ and $\operatorname{supp}\phi \subset B$. Then $\tilde\chi := \phi * \chi \in L^p(Z)^\infty$ with $\tilde\chi(a_{-k}z_0) \geq k$. Hence $\tilde\chi$ is unbounded.

In general, given a unipotent subalgebra $\mathfrak{n}$, there does not necessarily exist a semisimple element which normalizes $\mathfrak{n}$. For example if $U \in \mathfrak{g} = \mathfrak{sl}(5,\mathbb{C})$ is a principal nilpotent element, then $\mathfrak{n} = \operatorname{span}\{U, U^2 + U^3\}$ is a 2-dimensional abelian unipotent subalgebra which is not normalized by any semi-simple element of $\mathfrak{g}$. The next lemma offers a remedy out of this situation by finding an ideal $\mathfrak{n}_1 \lhd \mathfrak{n}$ which is normalized by a real semisimple element X with $\mathfrak{n} \subset \mathfrak{g}_X^+ + \mathfrak{g}_X^0$.

Lemma A.3. *Let $\mathfrak{n} \subset [\mathfrak{g},\mathfrak{g}]$ be an* ad*-nilpotent subalgebra and let $0 \neq U \in \mathfrak{z}(\mathfrak{n})$. Then there exists a real semi-simple element $X \in \mathfrak{g}$ such that $[X,U] = 2U$ and $\mathfrak{n} \subset \mathfrak{g}_X^+ + \mathfrak{g}_X^0$.*

Proof. According to the Jacobson-Morozov theorem one finds elements $X, V \in \mathfrak{g}$ such that $\{X, U, V\}$ form an $\mathfrak{sl}_2$-triple, i.e. satisfy the commutator relations $[X,U] = 2U$, $[X,V] = -2V$, $[U,V] = X$. Note that

$\mathfrak{n} \subset \mathfrak{z}_{\mathfrak{g}}(U)$ and that $\mathfrak{z}_{\mathfrak{g}}(U)$ is $\operatorname{ad} X$-stable. It is known and in fact easy to see that $\mathfrak{z}_{\mathfrak{g}}(U) \subset \mathfrak{g}_X^+ + \mathfrak{g}_X^0$. All assertions follow. □

Within the notation of Lemma A.3 we set $\mathfrak{n}_1 = \mathbb{R}U$ and $N_1 = \exp(\mathfrak{n}_1)$. Furthermore we set $Z_1 = G/N_1$ and consider the contractive averaging map

$$L^1(Z_1) \to L^1(Z), \quad f \mapsto \widehat{f}; \quad \widehat{f}(gN) = \int_{N/N_1} f(gnN_1)\, d(nN_1).$$

Let $B \subset G$ be a compact ball around $\mathbf{1}$, of sufficiently large size to be determined later, and let $B_1 = B \cdot B \subset G$. Let χ be the function on Z_1 constructed as in (A.1), using the element X from Lemma A.3 and the compact set B_1. Let $\widehat{\chi} \in L^1(Z)$ be the average of χ. We claim that $\widehat{\chi}(Ba_{-k}z_0) \geq k$ for all k. In fact let $Q \subset N/N_1$ be a compact neighborhood of $\mathbf{1}$ in N/N_1 with $\operatorname{vol}_{N/N_1}(Q) = 1$. Then for B large enough we have $a_{-k}Qa_k \subset B$ for all k (Lemma A.3). Hence for $b \in B$,

$$\widehat{\chi}(ba_{-k}z_0) \geq \int_Q \chi(ba_{-k}nN_1)\, d(nN_1) \geq k,$$

proving our claim.

To continue we conclude that $f_p := (\widehat{\chi})^{\frac{1}{p}} \in L^p(Z)$ is a function with $f_p(Ba_{-k}z_0) \geq k$ for all k. Finally we smoothen f_p as before and conclude that VAI does not hold true.

A.2. **The general case of a non-reductive unimodular space.** Finally we shall prove Proposition A.1 in the general situation where H is a non-reductive closed and connected subgroup for which $Z = G/H$ is unimodular.

Proof. We will argue by induction on $\dim \mathfrak{g}$. Suppose first that $\mathfrak{h}$ is contained in a proper subalgebra $\tilde{\mathfrak{h}}$ of $\mathfrak{g}$, which is reductive in $\mathfrak{g}$. Then $\mathfrak{h}$ is not reductive in $\tilde{\mathfrak{h}}$ ([13], §6.6 Cor. 2). By induction $\tilde{H}/H$ is not VAI, in the strong sense that for every $1 \leq p < \infty$ there exists an unbounded function $f \in L^p(\tilde{H}/H)^\infty$. We claim that G/H is not VAI in the same strong sense. Let $\tilde{\mathfrak{q}} \subset \mathfrak{g}$ be the orthogonal complement to $\tilde{\mathfrak{h}}$ in $\mathfrak{g}$. Then for a small neighborhood $V \subset \tilde{\mathfrak{q}}$ of 0 the tubular map

$$V \times \tilde{H} \to G, \quad (X, h) \mapsto \exp(X)h$$

is diffeomorphic. The Haar measure on G is expressed by $J(X)dXdh$ with $J > 0$ a bounded positive function. Since $\tilde{H}$ normalizes $\tilde{\mathfrak{q}}$, this allows us to extend smooth L^p-functions from $\tilde{H}/H$ to G/H and we see that G/H is not VAI in the strict sense.

We assume from now on that $\mathfrak{h}$ is not contained in any reductive proper subalgebra $\tilde{\mathfrak{h}}$.

To proceed we recall the characterization of maximal subalgebras of $\mathfrak{g}$ (see [14], Ch. 8, §10, Cor. 1). A maximal subalgebra is either a maximal parabolic subalgebra or it is a maximal reductive subalgebra. In the present case it follows that $\mathfrak{h}$ is contained in a maximal parabolic subalgebra $\mathfrak{p}_0$. We write $\mathfrak{p}_0 = \mathfrak{n}_0 \rtimes \mathfrak{l}_0$ where $\mathfrak{n}_0$ is the unipotent radical of $\mathfrak{p}_0$. Note that $\mathfrak{l}_0$ is reductive in $\mathfrak{g}$ and hence $\mathfrak{h}$ is not contained in $\mathfrak{l}_0$. In addition we may assume that $\mathfrak{s} \subset \mathfrak{l}_0$ ([13] §6.8 Cor. 1).

In the next step we claim that $\mathfrak{z}(\mathfrak{l}_0)$ is not contained in $\mathfrak{h}$. In fact as G/H is unimodular, $|\det \mathrm{Ad}(h)|_{\mathfrak{h}}| = 1$ for $h \in H$. If in addition $h \in Z(\mathfrak{l}_0)$, then h centralizes $\mathfrak{s}$ and it follows that $|\det \mathrm{Ad}(h)|_{\mathfrak{r}}| = 1$. Hence $\mathfrak{z}(\mathfrak{l}_0) \cap \mathfrak{h}$ centralizes $\mathfrak{r}$ as well. If $\mathfrak{z}(\mathfrak{l}_0)$ would be contained in $\mathfrak{h}$, then this would force $\mathfrak{h} \subset \mathfrak{l}_0$, since $\mathfrak{l}_0$ is the centralizer of its center. We would thus arrive at the already excluded case $\mathfrak{h} \subset \mathfrak{l}_0$.

We fix $0 \neq X \in \mathfrak{z}(\mathfrak{l}_0) \setminus \mathfrak{h}$ such that $\mathfrak{n}_0 \subset \mathfrak{g}_X^+$. As before we set $a_t := \exp(tX)$ and observe that $a_t z_0 \to \infty$ in Z for $|t| \to \infty$ (this is because $a_t[L_0, L_0]N_0$ goes to infinity in $G/[L_0, L_0]N_0$.)

Next we define a subspace $\mathfrak{l}_1 \subset \mathfrak{l}_0$ by

$$\mathfrak{l}_1 = \mathrm{pr}_{\mathfrak{l}_0}(\mathfrak{h})^\perp$$

where $\mathrm{pr}_{\mathfrak{l}_0} : \mathfrak{p}_0 \to \mathfrak{l}_0$ is the projection along $\mathfrak{n}_0$. Then $\mathfrak{l}_1 \oplus (\mathfrak{h} + \mathfrak{n}_0) = \mathfrak{p}_0$. In addition, we construct an $\mathrm{ad}\, X$-invariant subspace $\mathfrak{n}_1$ of $\mathfrak{n}_0$ such that $\mathfrak{h} + \mathfrak{n}_0 = \mathfrak{h} \oplus \mathfrak{n}_1$, as follows. If $\mathfrak{n}_0 \subset \mathfrak{h}$, then $\mathfrak{n}_1 = \{0\}$. Otherwise we choose an $\mathrm{ad}\, X$-eigenvector, say Y_1, in $\mathfrak{n}_0$ with largest possible eigenvalue, such that $\mathfrak{h} + \mathbb{R}Y_1$ is a direct sum. If this sum contains $\mathfrak{n}_0$, we set $\mathfrak{n}_1 = \mathbb{R}Y_1$. Otherwise we continue that procedure until a complementary subspace is reached. Now $\mathfrak{l}_1 \oplus \mathfrak{h} \oplus \mathfrak{n}_1 = \mathfrak{p}_0$.

Let us first exclude the case where $\mathfrak{n}_1 = \mathfrak{n}_0$, i.e. $\mathfrak{h} \cap \mathfrak{n}_0 = \{0\}$. In this situation the projection map $\mathrm{pr}_{\mathfrak{l}_0}|_{\mathfrak{h}} : \mathfrak{h} \to \mathfrak{l}_0$, which is a Lie-algebra homomorphism, is injective. Write $\mathfrak{h}_0$ for the homomorphic image of $\mathfrak{h}$ in $\mathfrak{l}_0$. The analysis will be separated in two cases, the first being that $\mathfrak{h}_0$ is reductive in $\mathfrak{l}_0$.

Assume $\mathfrak{h}_0$ is reductive in $\mathfrak{l}_0$. In particular it is a reductive Lie algebra, hence so is $\mathfrak{h}$. In the Levi decomposition $\mathfrak{h} = \mathfrak{s} \oplus \mathfrak{r}$ we now know that $\mathfrak{r}$ is the center of $\mathfrak{h}$. Let $\mathfrak{u}$ be the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{r}$ and $\theta(\mathfrak{r})$, then $\mathfrak{s} + \mathfrak{u}$ is a direct Lie algebra sum. Moreover, $\mathfrak{s} + \mathfrak{u}$ is θ-invariant, hence reductive in $\mathfrak{g}$, and hence in fact $= \mathfrak{g}$ by our previous assumption on $\mathfrak{h}$. Thus $\mathfrak{s}$ is an ideal in $\mathfrak{g}$ which we may as well assume is 0. Now $\mathfrak{h} = \mathfrak{r}$ is an abelian subalgebra which together with $\theta(\mathfrak{r})$ generates $\mathfrak{g}$. We shall reduce to the case where $\mathfrak{r}$ is nilpotent in $\mathfrak{g}$, which we already treated. Every element $X \in \mathfrak{r}$ has a Jordan decomposition $X_n + X_s$ (in $\mathfrak{g}$), and we let $\mathfrak{o}_1, \mathfrak{o}_2$ be the subalgebras generated by the X_n's and X_s's, respectively. Then $\mathfrak{o} = \mathfrak{o}_1 \oplus \mathfrak{o}_2$ is abelian and $\mathfrak{o}_2$

consists of semisimple elements. The centralizer of $\mathfrak{o}_2$ is reductive in $\mathfrak{g}$ and contains $\mathfrak{r}$, hence equal to $\mathfrak{g}$. Hence $\mathfrak{o}_2$ is central in $\mathfrak{g}$, and we may assume that it is θ-stable. Let $\mathfrak{g}_1$ be the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{o}_1$ and $\theta(\mathfrak{o}_1)$. It is reductive in $\mathfrak{g}$, and $(\mathfrak{g}_1, \mathfrak{o}_1)$ is of the type already treated, hence not VAI. Since $\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{o}_2$ we can now conclude that $(\mathfrak{g}, \mathfrak{r}) = (\mathfrak{g}, \mathfrak{h})$ is not VAI either.

We now assume that $\mathfrak{h}_0$ is not reductive in $\mathfrak{g}$. Let H_0 and L_0 be the connected subgroups of G corresponding to $\mathfrak{h}_0$ and $\mathfrak{l}_0$. As G/H is unimodular and H is homomorphic to H_0, it follows that G/H_0 and thus L_0/H_0 is unimodular. By induction we find for every $1 \leq p < \infty$ an unbounded function $f \in L^p(L_0/H_0)^\infty$. As before in the case of $\tilde{H}/H$ we extend f to a smooth vector in $L^p(G/H)$ (note that $P_0/H \to L_0/H_0$ is a fibre bundle, and we first extend f to a function on P_0/H and then to a function on G/H).

In the sequel we assume that $\mathfrak{n}_1$ is a proper subspace of $\mathfrak{n}_0$.

Let $\overline{\mathfrak{n}}_0$ be the nilradical of the parabolic opposite to $\mathfrak{p}_0$ and consider the $\operatorname{ad} X$-invariant vector space

$$\mathfrak{v} := \overline{\mathfrak{n}}_0 \times \mathfrak{l}_1 \times \mathfrak{n}_1 \subset \mathfrak{g}$$

which is complementary to $\mathfrak{h}$.

For fixed $t \in \mathbb{R}$ we consider the differentiable map

$$\begin{aligned} &\Phi = \Phi_t : \mathfrak{v} = \overline{\mathfrak{n}}_0 \times \mathfrak{l}_1 \times \mathfrak{n}_1 \to Z, \\ &(Y^-, Y^0, Y^+) \mapsto \exp(Y^-)\exp(Y^0)\exp(Y^+) a_t z_0 \,. \end{aligned}$$

With $y^\pm = \exp(Y^\pm)$ and likewise $y^0 = \exp(Y^0)$ we get for the differential of Φ:

$$\begin{aligned} d\Phi(Y^-, Y^0, Y^+)(X^-, X^0, X^+) &= d\tau_{y^- y^0 y^+ a_t}(z_0) \circ \operatorname{Ad}(a_t)^{-1} \circ \\ \circ\Big(\operatorname{Ad}(y_0 y^+)^{-1} \frac{\mathbf{1} - e^{-\operatorname{ad} Y^-}}{\operatorname{ad} Y^-}(X^-) &+ \operatorname{Ad}(y^+)^{-1} \frac{\mathbf{1} - e^{-\operatorname{ad} Y^0}}{\operatorname{ad} Y^0}(X^0) + \\ + \frac{\mathbf{1} - e^{-\operatorname{ad} Y^+}}{\operatorname{ad} Y^+}(X^+) + \mathfrak{h}\Big) \,. \end{aligned}$$

In order to estimate the Jacobian of $d\Phi$ we will identify $T_{y^- y^0 y^+ a_t z_0} Z$ with $\mathfrak{v}$ via the map

$$T_{y^- y^0 y^+ a_t z_0} Z \to \mathfrak{v}, \quad d\tau_{y^- y^0 y^+ a_t}(z_0)(X + \mathfrak{h}) \mapsto \pi_{\mathfrak{v}}(X + \mathfrak{h})$$

where $\pi_{\mathfrak{v}} : \mathfrak{g} \to \mathfrak{v}$ is the projection along $\mathfrak{h}$. Within this notation we obtain

$$d\Phi(Y^-, Y^0, Y^+)(X^-, X^0, X^+) = \pi_{\mathfrak{v}} \circ \operatorname{Ad}(a_t)^{-1} \circ \\ \circ\Big(\operatorname{Ad}(y_0 y^+)^{-1} \frac{\mathbf{1} - e^{-\operatorname{ad} Y^-}}{\operatorname{ad} Y^-}(X^-) + \operatorname{Ad}(y^+)^{-1} \frac{\mathbf{1} - e^{-\operatorname{ad} Y^0}}{\operatorname{ad} Y^0}(X^0) + \\ + \frac{\mathbf{1} - e^{-\operatorname{ad} Y^+}}{\operatorname{ad} Y^+}(X^+)\Big).$$

Let $Y = (Y^-, Y^0, Y^+) \in \mathfrak{v}$. It follows that there exists a linear map $L(Y) : \mathfrak{v} \to \mathfrak{g}$ such that

$$d\Phi_t(Y) = \pi_{\mathfrak{v}} \operatorname{Ad}(a_t)^{-1}(\mathbf{1}_{\mathfrak{v}} + L(Y))$$

for all t, and that $\|L(Y)\| \to 0$ for $Y \to 0$. We rewrite as

$$\text{(A.2)} \qquad d\Phi(Y) = \operatorname{Ad}(a_t)^{-1}(\mathbf{1}_{\mathfrak{v}} + \operatorname{Ad}(a_t)\pi_{\mathfrak{v}} \operatorname{Ad}(a_t)^{-1} L(Y))$$

In order to control the remainder term, we define

$$M_t := \sup_{U \in \mathfrak{g}, \|U\|=1} \| \operatorname{Ad}(a_t)\pi_{\mathfrak{v}} \operatorname{Ad}(a_t)^{-1} U\|$$

and claim that M_t stays bounded for $t \to -\infty$. For $U \in \mathfrak{v}$ we have $\operatorname{Ad}(a_t)\pi_{\mathfrak{v}} \operatorname{Ad}(a_t)^{-1} U = U$, hence we may assume $U \in \mathfrak{h}$. Since $\mathfrak{h} \subset \mathfrak{p}_0$ we can write U as a combination of an element $Y_0 \in \mathfrak{l}_0$ and possibly some $\operatorname{ad} X$-eigenvectors Y_λ with eigenvalues $\lambda > 0$. Then

$$\operatorname{Ad}(a_t)^{-1} U = Y_0 + \sum e^{-\lambda t} Y_\lambda = U + \sum (e^{-\lambda t} - 1) Y_\lambda$$

(possibly with an empty sum). If $Y_\lambda \in \mathfrak{n}_1$ then

$$\operatorname{Ad}(a_t)\pi_{\mathfrak{v}}(e^{-\lambda t} - 1) Y_\lambda = (1 - e^{\lambda t}) Y_\lambda \to Y_\lambda$$

as $t \to -\infty$. On the other hand it follows from the definition of $\mathfrak{n}_1$, that if Y_λ is not in $\mathfrak{n}_1$ then either it belongs to $\mathfrak{h}$ or it is a sum of an element from $\mathfrak{h}$ and some eigenvectors $V_\mu \in \mathfrak{n}_1$ with eigenvalues $\mu \geq \lambda$. Then

$$\operatorname{Ad}(a_t)\pi_{\mathfrak{v}}(e^{-\lambda t} - 1) Y_\lambda = \sum e^{\mu t}(e^{-\lambda t} - 1) V_\mu$$

(possibly with an empty sum), which stays bounded for $t \to -\infty$. Our claim is thus established, and it follows that $\operatorname{Ad}(a_t)^{-1}\mathbf{1}_{\mathfrak{v}}$ dominates (A.2), for $Y \in \mathfrak{v}$ sufficiently small.

Combining our reasoning, and using that $\mathfrak{n}_1$ is proper in $\mathfrak{n}_0$, we obtain for every small enough compact neighborhood $Q \subset \mathfrak{v}$ of 0 some constants $c_Q, C_Q > 0$ such that

$$c_Q e^{t\gamma} \leq \sup_{Y \in Q} |\det d\Phi_t(Y)| \leq C_Q e^{t\gamma} \qquad (t \leq 0)$$

for some $\gamma > 0$ independent of Q. In particular $\Phi_t|_Q$ is a chart.

Fix now such a compact neighborhood Q, and let $\psi \in C_c^\infty(Q)$ be a function with $0 \leq \chi \leq 1$ and $\psi(0) = 1$. For all $t < 0$ define $\chi_t \in C_c^\infty(Z)$ by

$$\chi_t(z) = \psi(\Phi_t^{-1}(z)) \qquad (z \in Z),$$

then $\chi_t \in L^p(Z)$ for all $1 \leq p < \infty$ and $t \leq 0$, with $\|\chi_t\|_p \leq Ce^{t\gamma/p}$ for some $C > 0$ not depending on t (but possibly on p). Finally we set

$$\chi := \sum_{n\in\mathbb{N}} n\chi_{-n}\,,$$

Then $\chi \in L^p(Z)$ for all $1 \leq p < \infty$. It is also clear that $\chi \in C^\infty(Z)$ and that χ is unbounded. It remains to be seen that $\chi \in L^p(Z)^\infty$.

Let $U \in \mathfrak{g}$ and consider the left derivatives $L(U)\chi_t$. At $z = \Phi_t(Y) = y^- y_0 y^+ a_t z_0$ these are given by

$$L(U)\chi_t(z) = d/ds|_{s=0}\, \chi_t(\exp(sU)z).$$

For Y in a compact set, we may as well consider the derivatives of

$$\chi_t(y^- y_0 y^+ \exp(sU)a_t z_0).$$

Notice that $\exp(sU)z_0 \simeq \exp(s\pi_{\mathfrak{v}}U)z_0$ for $s \to 0$ and rewrite the derivative as

$$d/ds|_{s=0}\, \chi_t(y^- y_0 y^+ \exp(s\,\mathrm{Ad}(a_t)^{-1}\pi_{\mathfrak{v}}\,\mathrm{Ad}(a_t)U)a_t z_0).$$

We conclude that this is a linear combination of derivatives of ψ on Q, with coefficients that depend smoothly on Y and are bounded for $t < 0$. Here the previously attained bound on M_t is used. As before we conclude $L(U)\chi_t \in L^p(Z)$ for all $t \leq 0$, with exponentially decaying p-norms. It follows that $L(U)\chi \in L^p(Z)$. By repeating the argument for higher derivatives we finally see that $\chi \in L^p(Z)^\infty$. □

Appendix B. Proof of Lemma 7.6

We assume that Z is algebraic, that $P = MAN$ is a parabolic subgroup such that PH is open, that $M/M \cap H$ is compact and that $L := P \cap H \subset M$. According to (7.4)

$$\text{(B.1)} \qquad \exists a \in A: \quad \mathfrak{a} \oplus (\mathfrak{k}^a + \mathfrak{h}) = \mathfrak{g}.$$

The statement in the lemma is that $G = KAH$.

We first claim that we may assume that PH is open for all *all* choices of parabolics with $A_P = A$. Note that both (B.1) and the openess of PH are open conditions. This implies that if we replace A by its conjugate by an element $k \in K$, then both conditions hold for generic k. It follows that in each neighborhood of $\mathbf{1}$ in K we can find an element k such that after replacing A by $k^{-1}Ak$, then (B.1) holds and PH is open for all $P \supset A$. But if we know that $KAkH = G$ for all k in a sequence $k \to \mathbf{1}$, then $KAH = G$. This proves the claim and allows us to make the asserted assumption. Notice however that conjugating P possibly destroys that $L \subset M$ for $L = P \cap H$. However, Lemma B.2 applied to $(R, S) = (P, H), (H, P)$ implies that H/L and P/L stay unimodular homogeneous spaces. We claim that $L \subset MN$. In fact if $P = P_k := kPk^{-1}$ and $L_k := P_k \cap H$ then Lemma B.1 implies that $L_k = (H \cap (M_k)^n)(A \cap (M_k)^n)(H \cap N_k)$. Note that $n \in N_k$ depends on k but $n \to \mathbf{1}$ for $k \to \mathbf{1}$. The Lie algebras $\mathfrak{l}_k$ converge to $\mathfrak{l} \subset \mathfrak{m}$, and thus $L_k = H \cap (M_k)^n \subset M_k N_k$ as was to be shown.

Note that $K_\mathbb{C} A_\mathbb{C} H_\mathbb{C}$ being the image of the algebraic morphism

$$K_\mathbb{C} \times A_\mathbb{C} \times H_\mathbb{C} \to G_\mathbb{C},\ (k, a, h) \mapsto kah$$

is an affine subvariety of $G_\mathbb{C}$ which is of maximal dimension as it is submersive at a point. It follows that $K_\mathbb{C} A_\mathbb{C} H_\mathbb{C}$ contains a Zariski open subset of $G_\mathbb{C}$ and $G \cap K_\mathbb{C} A_\mathbb{C} H_\mathbb{C}$ has dense interior in G.

Let $C_\mathbb{C} \subset (K_\mathbb{C} A_\mathbb{C} H_\mathbb{C})^o$ be a connected component (with respect to the Zariski topology). We consider the sets $C := C_\mathbb{C} \cap G$. Since these are real algebraic varieties, every component that intersects non-trivially with KAH is necessarily contained in KAH. Hence it suffices to prove that $C \cap KAH$ is non-empty for each component

$$C \subset G \cap (K_\mathbb{C} A_\mathbb{C} H_\mathbb{C})^o.$$

Observe first that C has unbounded $A_\mathbb{C}$-part, that is there exists a sequence $p_n = k_n x_n h_n \in C$ with $k_n \in K_\mathbb{C}$, $h_n \in H_\mathbb{C}$ and with $x_n \in A_\mathbb{C}$ tending to infinity. Otherwise the projection of C in the categorical quotient $Q_\mathbb{C} := G_\mathbb{C} // H_\mathbb{C} \times K_\mathbb{C}$ (see [32], Sect. II.3.2) would be bounded. Hence the image of C would be a bounded subvariety of $Q_\mathbb{R}$ with interior points, and this is not possible. To finish the proof, we shall show that eventually $p_n \in KAH$.

Let $T := \exp(i\mathfrak{a})$ and note that $A_\mathbb{C} = AT$. Further we set $A_\mathbb{R} = A_\mathbb{C} \cap G$ and note that $A_\mathbb{R} = AF$ for $F \subset T \cap K$ a finite 2-group. We write $x_n = a_n t_n$ with $a_n \in A$ and $t_n \in T$. Passing to a subsequence if necessary we may assume that a_n lies in the closure of a fixed chamber A^+. We claim that $t_n \in F$ for sufficiently large n. This is the crucial part of the proof.

It is here that we need analytic arguments. We shall provide a sufficient supply of analytic $K \times H$-invariant functions on G. We work with generalized principal series: each $\lambda \in \mathfrak{a}_\mathbb{C}^*$ defines a character $\chi_\lambda = 1 \otimes e^\lambda \otimes 1 : P \to \mathbb{C}^*$ and defines a line bundle $G \times_P \mathbb{C}_\lambda \to G/P$. The smooth sections of this line bundle we denote by E_λ. As a topological vector space we identify E_λ with $C^\infty(K/K_M)$ in a K-equivariant way (compact model). Here $K_M = K \cap M$.

To continue we recall that $L = H \cap P \subset MN$. We claim that

$$\eta_\lambda : E_\lambda \to \mathbb{C}, \quad f \mapsto \int_{H/L} f(h)\, dh$$

defines an invariant functional provided $\operatorname{Re}\lambda \in \mathfrak{a}^*$ is large and lies in $\mathfrak{a}^{*,+}$. Our concern is with convergence only. Recall that $G/P \simeq K/K_M$ and write $\mathcal{O} \subset K/K_M$ for the open subset corresponding to $HP/P \subset G/P$. In the sequel we use

$$H \times_L M \times A \times N \ni (h, m, a, n) \mapsto hman \in HP \subset G$$

as coordinates on the open set $HP \subset G$ and note that in $HP \ni g = h(g)m(g)a(g)n(g)$ the entry $a(g) \in A$ is defined whereas $h(g)$ is defined mod L. In particular, $kK_M \mapsto h(k)L$ is defined on $\mathcal{O}$. For a measurable function ϕ on H/L, the integration formula in Lemma B.3 gives

$$\int_{H/L} \phi(hL)\, d(hL) = \int_{\mathcal{O}} \phi(h(k))a(k)^{-2\rho}\, dk\,.$$

In particular, for $f \in E_\lambda$ it follows that

$$\int_{H/L} f(hL)\, d(hL) = \int_{\mathcal{O}} f(k)a(k)^{\lambda-2\rho}\, dk\,.$$

Let $P_0 = M_0A_0N_0 \subset P$ be a minimal parabolic, as in the proof of Lemma 6.3, such that P_0H is open. Recall the closed embedding (7.1). Let $\mu \in \mathfrak{a}_0^*$ be the highest weight of the corresponding finite dimensional representation π of G whose highest weight ray is fixed by P_0. Then the fact that P_0H is open allows the conclusion that $\langle v_\mu, v_H\rangle \neq 0$, and in particular that $\mu = 0$ on $\mathfrak{a}_0 \cap \mathfrak{h}$. Furthermore the identity

$$a(g)^\mu \langle v_\mu, v_H\rangle = \langle v_\mu, \pi(g^{-1})v_H\rangle$$

allows us to conclude that a^μ is continuous on HP.

Note that with μ_1 and μ_2 being highest weights of finite dimensional H-spherical representations, the same holds for $\mu_1 + \mu_2$ (take tensor products). Thus the H-spherical highest weights form a semi-group. By (B.1) we see that the categorical quotient $G_\mathbb{C}//H_\mathbb{C} \times K_\mathbb{C}$ has dimension equal to $\dim \mathfrak{a}$. It follows that the elements of the semi-group span $\mathfrak{a}^*$. Let us denote by $\mathbf{C} \subset \mathfrak{a}^*$ the convex hull of all $\mu's$. Hence η_λ is defined and non-zero for all $\lambda \in \mathbf{C}_\rho := 2\rho + \mathbf{C}$.

For $\lambda \in \mathbf{C}_\rho$ the matrix coefficient $\eta_\lambda(\pi(g)^{-1}v_K)$, where $v_K = \mathbf{1}_{K/M} \in E_\lambda^K$, is a smooth $K \times H$-invariant real function on G. We define

$$f_\lambda(g) := \eta_\lambda(\pi_\lambda(g)^{-1}v_K), \quad h_\lambda(g) := \operatorname{Re}[cf_\lambda(g)], \qquad (g \in G)$$

with $c = c(\lambda) \in \mathbb{C}$ a non-zero constant to be determined.

We assume first that $a_n \in A^+$, uniformly away from walls. Note that $\exp_{A_\mathbb{C}}^{-1}(\mathbf{1}) =: \Lambda$ defines a lattice in $\mathfrak{t}$ and that we can take $F = \exp(i\frac{1}{2}\Lambda)$. We let $\Omega \subset \mathfrak{t}$ be the interior of the standard fundamental domain for the action of $\frac{1}{2}\Lambda$ and set $T' := \exp(i\Omega)$ so that $D := A^+T' \subset A_\mathbb{C}$ is simply connected. As we are free to replace x_n by an appropriate representative in x_nF we may (after possibly deforming Ω slightly) in fact assume that $x_n \in D$.

Our proof of Th. 6.4 yields, as in [46], Sect. 4.4, an asymptotic expansion:

$$f_\lambda(a) \sim \sum_{w\in W}\sum_{\nu\in\mathbb{N}_0[\Sigma^+]} c_{w,\nu}a^{w\lambda-\rho-\nu} \quad (a \in A^+)\,.$$

In particular, $c_{\mathbf{1},0} \neq 0$, so that we can fix $c := c_{\mathbf{1},0}^{-1}$. Then

$$h_\lambda(a) \sim a^{\lambda-\rho} + \sum_{w\lambda-\rho-\nu\neq 0} \mathrm{Re}[cc_{w,\nu}]a^{w\lambda-\nu}.$$

The expansion converges for $a \in D$ and allows to define $h_\lambda(a)$ for $a \in D$. In these formulas we suppressed logarithmic terms which can be avoided under suitable regularity assumptions. We have thus extended h_λ to a real analytic function on $K_\mathbb{C}DH_\mathbb{C}$ which is $K_\mathbb{C} \times H_\mathbb{C}$-invariant. As $p_n \in G$ we obtain

$$\mathbb{R} \ni h_\lambda(p_n) = h_\lambda(a_nt_n) \simeq (a_nt_n)^\lambda$$

for n large and $\lambda \in \mathbf{C}_\rho$ generic. Thus t_n^λ is close to the real axis if n is sufficiently large. We claim that $t_n = \mathbf{1}$ provided n is sufficiently large. For that we write $t_n = \exp(iX^n)$ with $X^n \in \Omega$. The fact that t_n^λ is close to $\mathbb{R}$ means more precisely that for each $\epsilon > 0$ and each $\lambda \in \mathbf{C}_\rho$ there exists $N_\epsilon(\lambda) \in \mathbb{N}$ such that

$$\lambda(X^n) \in \pi\mathbb{Z} + (-\epsilon, \epsilon)\,.$$

for $n \geq N_\epsilon(\lambda)$. As Ω is bounded, we have in addition

$$|\lambda(X^n)| \leq C\|\lambda\|.$$

Let us fix a basis $\mu_1, \ldots, \mu_r$ of $\mathfrak{a}^*$ which is contained in $\mathbf{C}_\rho$ and introduce coordinates on Ω, say $X_j := \mu_j(X)$ for $X \in \Omega$. With regard to our basis we express $\lambda \in \mathbf{C}_\rho$ as $\lambda = \sum_{j=1}^r \lambda_j\mu_j$ with $\lambda_j \in \mathbb{R}$. Let $d > 1$ be an irrational number. Then by taking λ equal to each of the finitely many elements $\mu_1, \ldots, \mu_r, d\mu_1, \ldots, d\mu_r$ in $\mathbf{C}_\rho$ we obtain a constant $C > 0$ and for every $\epsilon > 0$ an $N_\epsilon \in \mathbb{N}$ such that

$$X_j^n \in [\pi\,\mathbb{Z} + (-\epsilon, \epsilon)] \cap [\frac{\pi}{d}\mathbb{Z} + (-\epsilon, \epsilon)] \cap [-C, C]$$

for all $n \geq N_\epsilon$ and $1 \leq j \leq r$. Clearly this set has only the element 0 provided $\epsilon > 0$ is small enough. This proves our claim.

The case where a_n does not lie in A^+, only in $\overline{A^+}$, is treated via the series regrouping argument in [16], Sect. 5, which allows control along the walls.

As $F \subset K$ we conclude that our component C contains a point $p = kah$ with $k \in K_\mathbb{C}$ and $h \in H_\mathbb{C}$ and $a \in A$. As C is left K-invariant, we may assume that $k = \exp(iY)$ for $Y \in \mathfrak{k}$ (use that $K_\mathbb{C} = K\exp(i\mathfrak{k})$).

Write $\tau : G_\mathbb{C} \to G_\mathbb{C}$ for the associated complex conjugation which has G as the set of real points. As $p \in G$ we get that $kah = \tau(kth) = k^{-1}a\tau(h)$ and thus

$$\exp(2i\,\mathrm{Ad}(a^{-1})Y) = a^{-1}k^2a = \tau(h)h^{-1} \in H_\mathbb{C}$$

As $2i\,\mathrm{Ad}(a^{-1})Y \in \mathfrak{g}_\mathbb{C}$ is a semi-simple element with real spectrum it follows that $i\,\mathrm{Ad}(a^{-1})Y \in \mathfrak{h}_\mathbb{C}$. Hence $p = kah = ah'h$ for some $h' \in H_\mathbb{C}$ and as p lies in G we get that $p \in AH$. This completes the proof. □

Lemma B.1. *Let H be an algebraic subgroup of G and $P < G$ a parabolic subgroup with Langlands decomposition $P = MAN$. Then there exists $n \in N$ such that*

$$H \cap P = (H \cap M^n)(H \cap A^n)(H \cap N)$$

where $M^n = nMn^{-1}$ and $A^n = nAn^{-1}$.

Proof. The algebraic group $S := H \cap P$ has a Levi decomposition $S = LU$ with U the unipotent radical. As N is the unique maximal unipotent subgroup of P it follows that $U \subset N$. Further, any Levi subgroup of P is conjugate under N to MA. As Levi subgroups of P coincide with maximal reductive subgroups of P, the assertion of the lemma follows. □

Lemma B.2. *Let G be a connected real linear algebraic group and R, S algebraic subgroups of G. Suppose that $RS \subset G$ is open and $R/R \cap S$ is unimodular. Then all open R-orbits in G/S are unimodular, i.e. if RxS is open for some $x \in G$, then $R/R \cap xSx^{-1}$ is unimodular.*

Proof. As RS is open, the same holds for $R_\mathbb{C}S_\mathbb{C}$ in $G_\mathbb{C}$. Note that $R_\mathbb{C}S_\mathbb{C}$ is Zariski open in $G_\mathbb{C}$. It follows that there exists only one open $R_\mathbb{C}$-orbit in $G_\mathbb{C}/S_\mathbb{C}$. This $R_\mathbb{C}$-orbit is unimodular and as any open R-orbit in G/S has the same complexification, the assertion follows. □

Lemma B.3. *Let $Z = G/H$ be of reductive type and (P, H) is a unimodular spherical pair such that $L := H \cap P$ is reductive in G. Then,* *up to normalization of measures,*

$$\int_{H/L} \phi(hL)\ d(hL) = \int_{K\cap HP} \phi(h(k))a(k)^{-2\rho}\ dk$$

for $\phi \in L^1(H/L)$*. In this formula one has* $HP \ni g = h(g)m(g)a(g)n(g)$ *according to* $P = MAN$*.*

Proof. Let ϕ be a test function on H/L. Choose a test function Φ on H such that

$$\phi(hL) = \Phi^L(hL) := \int_L \Phi(hl)\ dl\,.$$

Next choose a test function ψ on P with $\int_P \psi(p)\ d_rp = 1$ where d_rp is a right Haar measure on P. Define a function F on $H \times P$ by $F(h,p) = \Phi(h)\psi(p)$. We let L act on $H \times P$ as $(h,p) \mapsto (hl, l^{-1}p)$ and write $f := F^L$ for the corresponding fiber integral. Up to normalization of the Haar measure dg of G we then have

$$\int_{HP} f(g)\ dg = \int_{H/L} \phi(hL)\ d(hL)\,.$$

In fact

$$\begin{aligned}\int_{HP} f(g)dg &= \int_H \int_P F(h,p)\ d_rp\ dh\\ &= \int_H \int_P \Phi(h)\psi(p)\ d_rp\ dh\\ &= \int_H \Phi(h)\ dh = \int_{H/L} \phi(hL)\ d(hL)\,.\end{aligned}$$

On the other hand we have

$$\begin{aligned}\int_{HP} f(g)dg &= \int_K \int_P f(kp)\ d_rp\ dk\\ &= \int_K \int_P f(h(k)p(k)p)\ d_rp\ dk\\ &= \int_K \int_P f(h(k)p)\ a(k)^{-2\rho}\ d_rp\ dk\\ &= \int_K \int_L \int_P F(h(k)l, l^{-1}p)\ a(k)^{-2\rho}\ d_rp\ dl\ dk\\ &= \int_K \int_L \Phi(h(k)l)\ a(k)^{-2\rho}\ dl\ dk\\ &= \int_K \phi(h(k))\ a(k)^{-2\rho}\ dk\end{aligned}$$

where from line 4 to line 5 we use the fact that $a(l)^{-2\rho} = 1$ which holds true as $\mathrm{Ad}(L)$ acts measure preserving on P. $\square$

Appendix C. Proof of Lemma 11.5

Let $1 \leq p < \infty$ and suppose that $\Lambda \subset \widehat{G}_s$ is $L^p(Z)$-bounded. The assertion is that there exists $C > 0$ such that

$$\|m_{v,\eta}\|_p \leq C\|m_{v,\eta}\|_\infty \tag{C.1}$$

for all $\pi \in \Lambda$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ and $v \in \mathcal{H}_\pi^K$.

We recall the polar decomposition $G = K\overline{A_q^+}H$ of of Z (see Subsection 5.1.1). Let $\rho \in \mathfrak{a}_q^*$ be the associated Weyl half sum and recall that the Haar measure on Z, expressed in $K/M \times \mathfrak{a}_q^+$ coordinates is given by

$$J(X)\ dX\ d(kM) \tag{C.2}$$

with the Jacobian $J(X)$ satisfying the bound

$$J(X) \leq e^{2\rho(X)}\,. \tag{C.3}$$

It follows from [3] that a subset $\Lambda \subset \widehat{G}_s$ is $L^p(Z)$-bounded if and only if there exists a $\lambda \in \mathfrak{a}_q^*$ such that $\lambda\,|_{\overline{\mathfrak{a}_q^+}\setminus\{0\}} > 0$ and

- For all π, v and η as above there exists $C > 0$ such that
$$|m_{v,\eta}(a)| \leq Ca^\lambda \qquad (a \in \overline{A_q^+})\,.$$
- $\int_{\mathfrak{a}_q^+} e^{p\lambda(X)}J(X)\ dX < \infty$.

Denote in this proof by $\Sigma = \Sigma(\mathfrak{a}_q, \mathfrak{g}) \subset \mathfrak{a}_q^*$ the root system of $\mathfrak{a}_q$ and by $\mathcal{W}$ the corresponding Weyl-group. We embed $\mathfrak{a}_q$ in a maximal abelian subspace $\mathfrak{a} \subset \mathfrak{p}$, of which the corresponding Weyl group is then denoted by W. Let $\lambda_\pi \in \mathfrak{a}_\mathbb{C}^*$ be a spherical principal series parameter (unique up to W-conjugacy) of the representation π. Note that the real part of λ_π is uniformly bounded.

In the sequel we will say that λ_π is *generic* if the following two conditions are satisfied:

(1) $\langle \nu - 2w.\lambda_\pi, \nu\rangle \neq 0$ for all $\nu \in \mathbb{N}_0[\Sigma^+]\setminus\{0\}$ and $w \in W$.
(2) λ_π is regular, i.e. if $w.\lambda_\pi - \lambda_\pi \in \mathbb{Z}\Sigma^+$, then $w = \mathbf{1}$.

The first condition is related to the recurrence relation in [8], Prop. 5.2, and we note that it is automatically satisfied if λ_π is purely imaginary. The second condition is related to the linear independence of the basic functions Φ_{λ_π}.

In the sequel we let $f = m_{v,\eta}|_{A_q^+}$. If λ_π is generic then we have an asymptotic expansion of the form (see [5], Section 17):

$$f(a) = \sum_{w\in W}\sum_{\nu\in\mathbb{N}_0[\Sigma^+]} c_{w,\nu} a^{-\rho+w.\lambda_\pi-\nu} \qquad (a \in A_q^+)\,.$$

A similar expansion holds in the non-generic case, but with additional logarithmic terms of bounded degree.

We first work under the simplifying assumption that all elements of Λ are generic. In addition we will first assume that $\Lambda \subset i\mathfrak{a}^*$. It follows that there exists a constant $c > 0$ such that

$$|\langle \nu - 2w.\lambda_\pi, \nu\rangle| \geq c \tag{C.4}$$

for all $\pi \in \Lambda$, $\nu \in \mathbb{N}_0[\Sigma^+]\backslash\{0\}$ and $w \in W$. Then, in view of [7], Th. 7.4, it follows that there exist constants $C, \kappa > 0$, independent of $\pi \in \Lambda$, such that:

$$|c_{w,\nu}| \leq C(1 + \|\nu\|)^\kappa |c_{w,0}| \tag{C.5}$$

for all $w \in W$.

Let us call $Lf(a) := a^{-\rho}\sum_{w\in W} c_{w,0}a^{w\lambda_\pi}$ the leading term of f and put $\tilde{L}f(a) = a^\rho L(f)(a)$.

To normalize matters we request that

$$\max_{w\in W} |c_{w,0}| = 1\,. \tag{C.6}$$

Fix $X_0 \in \mathfrak{a}_q^+$ and choose a small compact ball U around 0 in $\mathfrak{a}_q$ such that $X_0 + U \subset \mathfrak{a}_q^+$. We let $C_U := \mathbb{R}^+U$ be the convex cone generated by U. For $R > 0$ we define a compact ball in $\overline{A_q^+}$ by

$$B_R := \exp(\{X \in C_U \mid \|X\| \leq R\})\,.$$

Further we let $S_R \subset B_R$ be the subset obtained by elements X with $\|X\| = R$.

To begin with we suppose that for all $t > 1$ there exists a constant $c_t > 0$ such that for all $R \geq 1$ one has

$$\|\tilde{L}(f)\|_{\infty, B_{tR}\backslash B_R} \geq c_t \tag{C.7}$$

for all $\pi \in \Lambda$ and all f subject to the normalization (C.6). Note that

$$\|f\|_\infty \geq \|f\|_{\infty,B_{tR}\backslash B_R} \geq \|L(f)\|_{\infty,B_{tR}\backslash B_R} - \|L(f) - f\|_{\infty,B_{tR}\backslash B_R}\,.$$

For all R we set $c_R := \min_{a\in B_R} a^{-\rho} = \min_{a\in S_R} a^{-\rho}$ and note that $c_R = e^{-\gamma_0 R}$ for some $\gamma_0 > 0$. Likewise we set $d_R := \max_{a\in S_R} a^{-\rho}$ and note that $d_R = e^{-\gamma_1 R}$. By choosing the "opening angle" U of C_U small enough we can make the difference $\gamma_0 - \gamma_1$ arbitrarily small.

Now observe that $\|L(f)\|_{\infty,B_{tR}\backslash B_R} \geq e^{-\gamma_0 tR}c_t$. Moreover, from (C.5) we obtain that f is approximated well by its leading term outside a ball of sufficiently large radius: there exists $\epsilon > 0$, $C > 0$ such that for sufficiently large R and all $t > 1$ one has

$$\|L(f) - f\|_{\infty, B_{tR}\setminus B_R} \leq Ce^{-(\gamma_1+\epsilon)R}$$

for all $\pi \in \Lambda$. Putting matters together, there exists a choice of C_U such that there exists a constant $c' > 0$ such that

$$\|f\|_\infty \geq c' \tag{C.8}$$

for all f subject to the normalization (C.6).

For any $R > 0$ observe that

$$\|f\|_p \leq \mathrm{vol}(B_R)^{\frac{1}{p}} \|f\|_{\infty, B_R} + \|f\|_{p, B_R^c}\,.$$

Here vol stands for the volume with respect to the measure $J(X)dX$. We claim, given $\delta > 0$ we find $R > 0$ such that $\|f\|_{p,B_R^c} < \delta$ uniformly. For $a \in A_q^+$ and the L^p-norm of functions restricted to $KaA_q^+.z_0$ this follows from (C.5). Along the walls we use the power series regrouping procedure from [16], Sect. 6, (adapted to the symmetric case in [3], Sect. 5) to conclude that the contributions along the walls stay uniformly small. Combining the claim with (C.8) we obtain a constant $c'' > 0$ such that

$$\|f\|_p \leq c''\|f\|_\infty\,.$$

This proves (C.1) subject to all the simplifying assumptions. Step by step we now drop those.

To begin with let us analyze what happens if (C.7) fails. It is instructive to see the rank one case, i.e. $\dim \mathfrak{a} = 1$. We choose coordinates x on $\mathfrak{a} = \mathbb{R}$ such that $a^{-\rho} = e^{-x}$ and note that

$$L(f)(x) = e^{-x}(ce^{\lambda x} + de^{-\lambda x})\,.$$

Here $\lambda = \lambda_\pi$ is a complex number with $|\operatorname{Re}\lambda| < 1$. We implement our normalization (C.6) by the request $c = 1$ and $|d| \leq 1$. We see that $(C.7)$ holds provided λ is bounded away from 0. The worst case scenario is for $\lambda = is$ small and imaginary and $d = -1$: then $L(f)(x) = 2ie^{-x}\sin(sx)$. In this case however we note that

$$f(x) = 2ie^{-x}\sin(sx) + e^{-x}\sum_{n=1}^\infty e^{-nx}(e^{isx}c_n + e^{-isx}d_n)\,.$$

The precise form of the recurrence relation in [7], Prop. 5.2, implies that

$$|e^{-nx}(e^{isx}c_n + e^{-isx}d_n)| \leq Ce^{(-n+1/2)x}|s|$$

for a uniform constant $C > 0$. Again $\|f\|_\infty$ and $\|f\|_p$ are controlled by the leading term as before and we are done.

Let us now analyze in general the case where (C.7) starts to fail. From the Lemma below we conclude that (C.7) holds true provided there exists an $\epsilon > 0$ such that

$$\inf_{w \in W \setminus \{\mathbf{1}\}} \|w.\lambda_\pi - \lambda_\pi\| \geq \epsilon$$

for all $\pi \in \Lambda$. Suppose now that this conditions starts to fail for some $w \in W \setminus \{\mathbf{1}\}$. If w is the longest element, then this means that $\lambda_\pi \to 0$ and, in view of the explicit recurrence relation in [7], we can argue as in the rank one case above. In general let us write $\Pi \subset \Sigma^+$ for the set of simple roots. Then we find a non-empty subset $F \subset \Pi$ such that $\langle \lambda_\pi, \alpha \rangle \to 0$ for $\alpha \in F$ whereas $|\langle \lambda_\pi, \alpha \rangle| \geq \epsilon > 0$ for all $\alpha \in \Pi \setminus F$. This reduces in essence the situation to a lower rank situation, i.e. in the limit we group $\sum_{w \in W/W_F} \tilde{c}_w a^{w.\lambda_\pi}$ with $\tilde{c}_w = \sum_{s \in W_F} c_{ws}$ and normalize with regard $\max_{w \in W/W_F} |\tilde{c}_w| = 1$.

We assumed that $\Lambda \subset i\mathfrak{a}^*$ and wish to drop this assumption. In fact the weaker condition (C.4) is enough and we analyze now when this starts to fail. Again it is instructive to understand the rank one case near a point where (1) fails. Say (1) fails near $\nu = n\alpha$, $n \in \mathbb{N}$ and α the simple root. Instead of using the leading term $L(f)$ we use now the n-th term to normalize matters, i.e.

$$L_n(f)(x) = e^{-nx}(c_n e^{i\lambda x} + d_n e^{-i\lambda x}) \, .$$

We then assume $\max\{|c_n|, |d_n|\} = 1$ and proceed as before. The higher rank case works in a similar way.

Finally we drop the genericity assumptions (1) and (2). As we said logarithmic terms appear now in the expansion. But since the bound holds for the dense set if generic parameters we obtain the bound for all parameters by continuity. □

Lemma C.1. *Let* $c = (c_1, \ldots, c_n) \in \mathbb{C}^n$, $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ *and* $\delta > 0$. *For the the function*

$$f_{\lambda,c}(t) := \sum_{j=1}^n c_j e^{i\lambda_j t} \qquad (t \in \mathbb{R}) \, .$$

the following inequality holds true:

$$[\sup_{|t| \leq \delta} |f_{\lambda,c}(t)|]^2 \geq \|c\|_2^2 - n(n-1)/2 \|c\|_\infty^2 \sup_{j \neq k} \frac{1}{\delta |\lambda_j - \lambda_k|} \, .$$

Proof. From

$$\int_{-\delta}^{\delta} |f_{\lambda,c}(t)|^2 \; dt \leq 2\delta [\sup_{|t| \leq \delta} |f_{\lambda,c}(t)|]^2$$

and

$$\int_{-\delta}^{\delta} |f_{\lambda,c}(t)|^2 \, dt = \sum_{j,k=1}^{n} c_j \overline{c_k} 2 \frac{\sin(\lambda_j - \lambda_k)\delta}{\lambda_j - \lambda_k}$$

the assertion follows. □

Appendix D. Concluding Remarks

D.1. **Remarks on VAI.** We investigated (VAI) with regard to unimodular homogeneous spaces $Z = G/H$ for which G was reductive and $H < G$ was a closed subgroup with finitely many connected components.

We did not address here the cases where H is not connected or G is not reductive. Without any further assumption let us assume that G is a connected Lie group and $H \subset G$ is a closed subgroup such that $Z = G/H$ is unimodular. In case G is infinitesimally simple and Z is not compact one might suspect that Z has VAI if and only if the Zariski closure of H is a proper reductive subgroup. For G and H algebraic and G reductive, H is reductive in G if and only if it is reductive. One might then suspect for G and H algebraic and G general, that Z has VAI if and only if the nilradical of H is contained in the nilradical of G.

Initially we wanted to prove the converse implication in Theorem 2.2 via a temperedness result for invariant measures. To be more specific, assume G and $H < G$ to be algebraic groups and $Z = G/H$ to be unimodular and quasi-affine. Under these assumptions we conjecture that there is a rational G-module V, and an embedding $Z \to V$ such that the invariant measure μ_Z, via pull-back, defines a tempered distribution on V. Note that if Z is of reductive type, then there exists a V such that the image of $Z \to V$ is closed, and hence μ_Z defines a tempered distribution V. If Z is not of reductive type, then all images $Z \to V$ are non-closed and our conjecture would imply that VAI does not hold. Our conjecture is supported by a result of Deligne, established in [42], which asserts that for a reductive group G and $X \in \mathfrak{g} := \mathrm{Lie}(G)$ the invariant measure on the adjoint orbit $Z := \mathrm{Ad}(G)(X) \subset \mathfrak{g}$ defines a tempered distribution on $\mathfrak{g}$. Various particular results in the theory of prehomogeneous vector spaces provide additional support for our conjecture (see [12]).

D.2. **Spaces of spherical type, polar type and property** (I)**.** In the context of polar type and Property (I) we state the following conjecture. All the conditions below are known to hold in case Z is symmetric, and also in the (non-reductive) case $H = N$ of Iwasawa decomposition. In case Z is algebraic we showed (Theorem 7.9 and

Theorem 8.5) that (3) $\Rightarrow$ (1)$\wedge$(2) under the additional hypothesis of unimodular spherical type.

Conjecture D.1. *Let G be linear simple group with finite center.*
(a) *For a unimodular homogeneous space $Z = G/H$ we expect the following properties to be equivalent:*

(1) *Z has property* (I)*;*
(2) *Z is of polar type;*
(3) *Z is of spherical type;*
(4) *For every irreducible admissible moderate growth smooth Fréchet representation (π, E) of G, the space of H-invariant functionals is of finite dimension.*

If in addition G is split, then the above should be equivalent to:

(5) *$Z_\mathbb{C} = G_\mathbb{C}/H_\mathbb{C}$ is a spherical variety (*[15]*).*

(b) *Let $\mathcal{Z}(\mathfrak{g})$ denote the center of the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$. If Z is of polar type with decomposition $G = KAH$, then all K-finite $\mathcal{Z}(\mathfrak{g})$-eigenfunctions on Z admit asymptotic expansions along A.*

D.3. **Main term counting for non-compact Y.** When $Y = G/\Gamma$ is non-compact, the Fourier inversion (10.5) involves continuous spectrum and it takes more care to define an appropriate space $\mathcal{A}(Y)$. One possibility is to use pseudo-Eisenstein series as in [20] or simply take $C_c(Y)_o^K$.

So let $f \in \mathcal{A}(Y)$ and

$$f = \int_{\widehat{G}_s} f_\pi \, d\mu(\pi)$$

its Plancherel-decomposition: here $f_\pi = m_{v_\pi, f^\wedge(\pi)}$ as in (10.5) before. In the first step we would like to define f_π^H. In case the cycle $H/\Gamma_H \subset Y$ is compact, this causes no difficulties. In general, this goes under the term "regularization (of Eisenstein series)" which was established in a variety of cases (see [35] for some results). However, we would like to point out that all the technical tools (Paley-Wiener-theory) are available to establish regularization, but this will not be subject of this paper.

After we have defined f_π^H, main term analysis needs a more quantitative version of (I). To be more precise one needs to exhibit a $1 \le p < \infty$ and establish uniform quantitative control of $\|m_{v,\eta}^H\|_p$ for all $\pi \in \widehat{G}_s \backslash \{\mathbf{1}\}$ which appear in in $L^2(Y)^K$. In case Z is symmetric harmonic analysis is sufficiently developed to cope with this problem (see the key- Lemma 11.5 in the Section 11). For more general spaces, one needs to study asymptotic expansion in sufficient detail.

References

[1] J.F. Adams, "Lectures on Exceptional Lie Groups", Chicago Lecture in Mathematics, 1996

[2] D. Akhiezer, "Lie Group Actions in Complex Analysis", Aspects of Mathematics, Vieweg 1995

[3] E. van den Ban, *Asymptotic behaviour of matrix coefficients related to reductive symmetric spaces*, Indag. Math. **49** (1987), 225–249.

[4] E. van den Ban, *The principal series for a reductive symmetric space I. H-fixed distribution vectors*, Ann, Sci. Éc. Norm. Sup. **21** (1988), 359–412.

[5] E. van den Ban and H. Schlichtkrull, *Asymptotic expansions and boundary values of eigenfunctions on Riemannian symmetric spaces*, J. Reine Angew. math. **380** (1987), 108-165.

[6] E. van den Ban and H. Schlichtkrull, *Asymptotic expansions on symmetric spaces*, Harmonic analysis on reductive groups (Bowdoin 1989), 7987, Progr. Math., 101, Birkhuser, 1991.

[7] E. van den Ban and H. Schlichtkrull, *Fourier transform on a semisimple symmetric space*, Invent. Math. **130** (1997), no. **3**, 517–574.

[8] E. van den Ban and H. Schlichtkrull, *Expansions of Eisenstein integrals on semisimple symmetric spaces*, Ark. Mat. **35** (1997), 59–86.

[9] Y. Benoist and H. Oh, *Effective equidistribution of S-integral points on symmetric varieties*, preprint

[10] J. Bernstein and B. Krötz, *Smooth Fréchet globalizations of Harish-Chandra modules*, submitted

[11] J. Bernstein and A. Reznikov, *Estimates of Automorphic functions*, Moscow Math. J. **4 (1)** (2004), 19–37

[12] N. Bopp and H. Rubenthaler, *Local zeta functions attached to the minimal spherical series for a class of symmetric spaces*, Mem. Amer. Math. Soc. **174** (2005), no. 821.

[13] N. Bourbaki, *Groupes et algèbres de Lie, Chapitre 1*, Hermann 1960.

[14] N. Bourbaki, *Groupes et algèbres de Lie, Chapitre 7-8*, Hermann 1975.

[15] M. Brion, *Classification des espaces homogènes sphériques*, Compositio Math. **63** (1987), 189-208.

[16] W. Casselman and D. Milicic, *Asymptotic Behaviour of Matrix Coefficients of Admissible Representations*, Duke Math. J. **49 (4)**, 869–930

[17] J.-L. Clerc, T. Kobayashi, B. Ørsted and M. Pevzner, Generalized Bernstein–Reznikov integrals. Math. Annalen **349** (2011), 395 - 431.

[18] M. Cowling, *Sur les coefficients des representations unitaires des groupes de Lie simples*, Lecture Notes in Mathematics **739** (1979), 132–178.

[19] P. Delorme and V. Sécherre, *An analogue of the Cartan decomposition for p-adic symmetric spaces of split reductive p-adic groups*, preprint 2008.

[20] W. Duke, Z. Rudnick and P. Sarnak, *Density of integer points on affine homogeneous varieties*, Duke Math. J. **71** (1993), no. **1**, 143–179.

[21] M. Einsiedler, G. Margulis and A. Venkatesh, *Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces*, Invent. Math. **177** (2009), no. **1**, 137–212.

[22] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. **71** (1993), no. **1**, 181–209.

[23] A. Eskin, S. Mozes and N. Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2) **143** (1996), 253–299.

[24] M. Flensted-Jensen, Analysis on non-Riemannian symmetric spaces, Conf. Board Math. Sci. Series **61**, Amer. Math. Soc. 1986.

[25] A. Gorodnik and A. Nevo *Counting lattice points* , to appear in J. Reine Angew. Math.

[26] B.H. Gross and D. Prasad, *On the decomposition of a representation of* SO_n *when restricted to* SO_{n-1}, Canad. J. Math. **44** (1992), 974–1002.

[27] Harish-Chandra, Automorphic Forms on Semisimple Lie Groups, Springer LNM 62.

[28] S. Helgason, "Differential geometry, Lie groups, and symmetric spaces", Academic Press 1978.

[29] L. Ji, *The Weyl upper bound on the discrete spectrum of locally symmetric spaces*, J. Diff. Geom. **51 (1)** (1999), 97–147.

[30] D. Kelmer and P. Sarnak, *Strong spectral gaps for compact quotients of products of* $PSL(2,\mathbb{R})$, J. Eur. Math. Soc. (JEMS) **11** (2009), no. **2**, 283–313.

[31] T. Kobayashi, *Invariant measures on homogeneous manifolds of reductive type*, J. Reine Angew. Math., **490** (1997), 37–53.

[32] H. Kraft, "Geometrische Methoden in der Invariantentheorie", Vieweg 1984

[33] M. Krämer, *Sphärische Untergruppen in kompakten zusammenhängenden Gruppen*, Compositio Math. **38 (2)** (1979), 129–153.

[34] B. Krötz and H. Schlichtkrull, *On function spaces on symmetric spaces*, to appear

[35] E. Lapid and J. Rogawski, *Periods of Eisenstein series*, C.R. Acad. Sci. Paris **333** (2001) , Série 1, 513–516.

[36] N. Lohoué and S. Mustapha, *Sur les transformées de Riesz sur les espaces homogènes des groupes de Lie semi-simples*, Bull. Soc. Math. France **128 (4)** (2000), 485–495.

[37] G. D. Mostow, *Some new decomposition theorems for semi-simple groups*, Mem. Amer. Math. Soc. **14** (1955), 31–54.

[38] A.L. Onishchik and E.B. Vinberg, *Lie Groups and Lie Algebras III*, Encycl. Math. Sciencies **41**, Springer, 1991.

[39] R. Phillips and Z. Rudnick, *The circle problem in the hyperbolic plane*, J. Funct. Anal. **121** (1994), 78–116.

[40] N. S. Poulsen, *On* C^∞*-vectors and intertwining bilinear forms for representations of Lie groups*, J. Funct. Anal. **9** (1972), 87–120.

[41] D. Prasad, *Trilinear forms for representations of* $\mathrm{GL}(2)$ *and local* ϵ*-factors.* Compositio Math. **75** (1990), 1–46.

[42] Ranga Rao, *Orbital integrals in reductive groups*, Ann. of Math. **(2) 96** (1972), 505–510.

[43] Z. Rudnick and H. Schlichtkrull, *Decay of eigenfunctions on semisimple symmetric spaces*, Duke Math. J. **64** (1991), 445-450.

[44] P. Sarnak, *Diophantine problems and linear groups*, Proceedings of the ICM 1990, Vol. I, 459–471, Math. Soc. Japan 1990.

[45] H. Schlichtkrull, *Hyperfunctions and harmonic analysis on symmetric spaces*, Birkhäuser 1984.

[46] N. Wallach, "Real Reductive Groups I", Academic Press, 1988

[47] J. Wolf, *Finiteness of orbit structure for real flag manifolds*, Geom. Dedicata **3** (1974), 377–384.