# ON QUANTUM UNIQUE ERGODICITY FOR LOCALLY SYMMETRIC SPACES I

LIOR SILBERMAN AND AKSHAY VENKATESH

ABSTRACT. We construct an equivariant microlocal lift for locally symmetric spaces. In other words, we demonstrate how to lift, in a "semi-canonical" fashion, limits of eigenfunction measures on locally symmetric spaces to Cartan-invariant measures on an appropriate bundle. The construction uses elementary features of the representation theory of semisimple real Lie groups, and can be considered a generalization of Zelditch's results from the upper half-plane to all locally symmetric spaces of noncompact type. This will be applied in a sequel to settle a version of the quantum unique ergodicity problem on certain locally symmetric spaces.

## 1. INTRODUCTION

1.1. **General starting point: the semi-classical limit on Riemannian manifolds.** Let $Y$ be a compact Riemannian manifold, with the associated Laplace operator $\Delta$ and Riemannian measure $d\rho$. An important problem of harmonic analysis (or mathematical physics) on $Y$ is understanding the behaviour of eigenfunctions of $\Delta$ in the large eigenvalue limit. The equidistribution problem asks whether for an eigenfunction $\psi$ with a large eigenvalue $\lambda$, $|\psi(x)|$ is approximately constant on $Y$. This can be approached "pointwise" and "on average" (bounding $\|\psi\|_{L^\infty(Y)}$ and $\|\psi\|_{L^p(Y)}$ in terms of $\lambda$, respectively), or "weakly": asking whether as $|\lambda| \to \infty$, the probability measures defined by $d\bar{\mu}_\psi(x) = |\psi(x)|^2 \, d\rho(x)$ converge in the weak-* sense to the "uniform" measure $\frac{d\rho}{\mathrm{vol}(Y)}$. For example, Sogge [19] derives $L^p$ bounds for $2 \leq p \leq \infty$, and in the special case of Hecke eigenfunctions on hyperbolic surfaces, Iwaniec and Sarnak [11] gave a non-trivial $L^\infty$ bound. Here we will consider the weak-* equidistribution problem for a special class of manifolds and eigenfunctions.

---

*Date*: 7th September 2004.

2000 *Mathematics Subject Classification.* Primary 81Q50; Secondary 11F, 37A45, 37D40, 22E45, 35P20.

The second author was supported in part by NSF Grant DMS-0245606. Part of this work was performed at the Clay Institute Mathematics Summer School in Toronto.

A general approach to the weak-* equidistribution problem was found by Šnirel′man [18]. To an eigenfunction $\psi$ he associates a distribution $\mu_\psi$ on the unit cotangent bundle $S^*Y$ projecting to $\bar{\mu}_\psi$ on $Y$. This construction (the "microlocal lift") proceeds using the theory of pseudo-differential operators and has the property that, for any sequence $\{\psi_n\}_{n=1}^{\infty} \subset L^2(Y)$ with eigenvalues $\lambda_n$ tending to infinity, any weak-* limit of the $\mu_n = \mu_{\psi_n}$ is a probability measure on the unit tangent bundle $S^*Y$, invariant under the geodesic flow. Since any weak-* limit of the $\mu_n$ projects to a weak-* limit of the $\bar{\mu}_n$, it suffices to understand these limits; Liouville's measure $d\lambda$ on $S^*Y$ plays here the role of the Riemannian measure on $Y$.

This construction has a natural interpretation from the point of view of semi-classical physics. The geodesic flow on $Y$ describes the motion of a free particle ("billiard ball"). $S^*Y$ is (essentially) the *phase space* of this system, i.e. the state space of the motion. In this setting one calls a function $g \in C^\infty(S^*Y)$ an *observable*. The state space of the quantum-mechanical billiard is $L^2(Y)$, with the infinitesimal generator of time evolution $-\Delta$. "Observables" here are bounded self-adjoint operators $B\colon L^2(Y) \to L^2(Y)$. Decomposing a state $\psi \in L^2(Y)$ w.r.t. the spectral measure of $B$ gives a probability measure on the spectrum of $B$ (which is the set of possible "outcomes" of the measurement). The expectation value of the "measuring $B$ while the system is in the state $\psi$" is then given by the matrix element $\langle B\psi, \psi\rangle$. In the particular case where $B$ is a pseudo-differential operator with symbol $g \in C^\infty(S^*Y)$, we think of $B$ as a "quantization" of $g$, and any such a $B$ will be denoted $\mathrm{Op}(g)$.

We can now describe Šnirel′man construction: it is given by $\mu_\psi(g) = \langle \mathrm{Op}(g)\psi, \psi\rangle$. This indeed lifts $\bar{\mu}_\psi$, since for $g \in C^\infty(Y)$ we can take $\mathrm{Op}(g)$ to be multiplication by $g$. If the $\psi$ are taken to be eigenfunctions then, asymptotically, this construction does not depend on the choice of "quantization scheme," that is to say, on the choice of the assignment $g \mapsto \mathrm{Op}(g)$. Indeed, if $B_1, B_2$ have the same symbol of order 0, and $-\Delta\psi = \lambda\psi$ (i.e. "$\psi$ is an eigenstate of energy $\lambda$") then one has $\langle (B_1 - B_2)\psi, \psi\rangle = O(\lambda^{-1/2})$.

On a philosophical level we expect that at the limit of large energies, our quantum-mechanical description to approach the classical one. We will not formalize this idea (the "correspondence principle"), but depend on it for motivating our main question, whether ergodic properties of the classical system persist in the semi-classical limit of the "quantized" version:

**Problem 1.1.** (Quantum Ergodicity) Let $\{\psi_n\}_{n=1}^{\infty} \subset L^2(Y)$ be an orthonormal basis consisting of eigenfunctions of the Laplacian.

(1) What measures occur as weak-* limits of the $\{\bar{\mu}_n\}$? In particular, when does $\bar{\mu}_n \xrightarrow[n\to\infty]{\text{wk-*}} d\rho$ hold?

(2) What measures occur as weak-* limits of the $\{\mu_n\}$? In particular, when does $\mu_n \xrightarrow[n\to\infty]{\text{wk-*}} d\lambda$ hold?

**Definition 1.2.** *Call a measure $\mu$ on $S^*Y$ a (microlocal)* quantum limit *if it is a weak-* limit of a sequence of distributions $\mu_{\psi_n}$ associated, via the microlocal lift, to a sequence of eigenfunctions $\psi_n$ with $|\lambda_n| \to \infty$.*

In this language, the main problem is classifying the quantum limits of the classical system, perhaps showing that the Liouville measure is the unique quantum limit. As formalized by Zelditch [23] (for surfaces of constant negative curvature) and Colin de Verdière [2] (for general $Y$), the best general result known is still:

**Theorem 1.3.** *Let $Y$ be a compact manifold, $\{\psi_n\}_{n=1}^{\infty} \subset L^2(Y)$ an orthonormal basis of eigenfunctions of $\Delta$, ordered by increasing eigenvalue. Then:*

(1) (Weyl's law; see e.g. [9]) $\frac{1}{N}\sum_{n=1}^{N}\mu_n \xrightarrow[N\to\infty]{\text{wk-*}} d\lambda$ holds with no further assumptions.

(2) (Šnirel′man-Zelditch-Colin de Verdière) Under the additional assumption that the geodesic flow on $S^*Y$ is ergodic, there exists a subsequence $\{n_k\}_{k=1}^{\infty}$ of density $1$ s.t. $\mu_{n_k} \xrightarrow[k\to\infty]{\text{wk-*}} d\lambda$.

**Corollary.** *For this subsequence, $\bar{\mu}_{n_k} \xrightarrow[k\to\infty]{\text{wk-*}} d\rho$.*

It was proved by Hopf [8] that the geodesic flow on a manifold of negative sectional curvature is ergodic. In this case, Rudnick and Sarnak [17] conjecture a simple situation:

**Conjecture 1.4.** *(Quantum unique ergodicity) Let $Y$ be a compact manifold of strictly negative sectional curvature. Then:*

(1) *(QUE on $Y$) $\bar{\mu}_n$ converge weak-* to the Riemannian measure on $Y$.*
(2) *(QUE on $S^*Y$) $d\lambda$ is the unique quantum limit on $Y$.*

We remark that [17] also gives an example of a hyperbolic $3$-manifold $Y$, a point $P \in Y$, and a sequence of eigenfunctions $\psi_n$ with eigenvalues $\lambda_n$, such that $|\psi_n(P)| \gg \lambda_n^{1/4-\epsilon}$. The point $P$ is a fixed point of many Hecke operators, and behaves in a similar fashion to the poles of a surface of revolution. This remarkable phenomenon does not seem to contradict Conjecture 1.4. In the sequel to this paper the scarcity of such points and their higher-dimensional analogues will play an important role.

1.2. **Past work: Quantum unique ergodicity on hyperbolic surfaces and $3$-manifolds.** The quantum unique ergodicity question for hyperbolic surfaces has been intensely investigated over the last two decades. We recall some important results.

Zelditch's work [22, 24] on the case of compact surfaces $Y$ of constant negative curvature provided a representation-theoretic alternative to the original construction of the microlocal lift via the theory of pseudo-differential operators. It is well-known that the universal cover of such a surface $Y$ is the upper half-plane $\mathbb{H} \simeq \mathrm{PSL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$, so $Y = \Gamma\backslash\mathbb{H}$ for a uniform lattice $\Gamma < G = \mathrm{PSL}_2(\mathbb{R})$. Then the $\mathrm{SO}_2(\mathbb{R}) \simeq S^1$ bundle $X = \Gamma\backslash\mathrm{PSL}_2(\mathbb{R}) \twoheadrightarrow Y$ is isomorphic to the unit cotangent bundle of $Y$. In this parametrization, the geodesic flow on $S^*Y$ is given by the action of the maximal split torus $A = \left\{ \begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix} \right\}$ on $X$ from the right. Zelditch's explicit microlocal lift starts with the observation that an eigenfunction $\psi_n$ (considered as a $K$-invariant function on $X$) can be thought of as the spherical vector $\varphi_0^{(n)}$ in an irreducible $G$-subrepresentation of $L^2(X)$. He then constructs another ("generalized") vector in this subrepresentation, a distribution $\delta^{(n)}$, and shows that the distribution given by $\mu_{\psi_n}(g) = \delta^{(n)}(g\overline{\varphi_0^{(n)}})$ for $g \in C_c^\infty(X)$ agrees (up to terms which decay as the $\lambda_n$ grow) with the microlocal lift. He then observes that the distribution $\mu_{\psi_n}$ is exactly annihilated by a differential operator of the form $H + \frac{J}{r_n}$ where $H$ is the infinitesimal generator of the geodesic flow, $J$ a certain (fixed) second-order differential operator, and $\lambda_n = -\frac{1}{4} - r_n^2$. It is then clear that any weak-* limit taken as $|\lambda_n| \to \infty$ will be annihilated (in the sense of distributions) by the differential operator $H$, or in other words be invariant under the geodesic flow. Wolpert [21] made Zelditch's approach self-contained by showing that the limits are positive measures without using pseudo-differential calculus. For a clear exposition of the Zelditch-Wolpert microlocal lift see [13].

Lindenstrauss's paper [13] considers the case of $Y = \Gamma\backslash(\mathbb{H} \times \cdots \times \mathbb{H})$ for an irreducible lattice $\Gamma$ in $\mathrm{PSL}_2(\mathbb{R}) \times \cdots \times \mathrm{PSL}_2(\mathbb{R})$. The natural candidates for $\psi_n$ here are not eigenfunctions of the Laplacian alone, but rather of all the "partial" Laplacians associated to each factor separately. Set now $G = \mathrm{PSL}_2(\mathbb{R})^h$, $K = \mathrm{SO}_2(\mathbb{R})^h$, $X = \Gamma\backslash G$, $Y = \Gamma\backslash G/K$, and take $\Delta_i$ to be the Laplacian operator associated with the $i$th factor (so that $\mathbb{C}[\Delta_1, \ldots, \Delta_h]$ is the ring of $K$-bi-invariant differential operators on $G$). Assume that $\Delta_i\psi_n + \lambda_{n,i}\psi_n = 0$, where $\lim_{n\to\infty} \lambda_{n,i} = \infty$ for each $1 \le i \le h$ separately. Generalizing the Zelditch-Wolpert construction, Lindenstrauss obtains distributions $\delta^{(n)}\overline{\varphi_0^{(n)}}$ on $X$, projecting to $\bar{\mu}_{\psi_n}$on $Y$, and so that every weak-* limit of these (a "quantum limit") is a finite positive measure invariant under the action of the full maximal split torus $A^h$. He then proposes the following version of QUE, also due to Sarnak:

**Problem 1.5.** (QUE on locally symmetric spaces) Let $G$ be a connected semi-simple Lie group with finite center. Let $K < G$ be a maximal compact subgroup, $\Gamma < G$ a lattice, $X = \Gamma\backslash G$, $Y = \Gamma\backslash G/K$. Let $\{\psi_n\}_{n=1}^{\infty} \subset L^2(Y)$ be a sequence of normalized eigenfunctions of the ring of $G$-invariant differential operators on $G/K$, with the eigenvalues w.r.t. the Casimir operator tending to $\infty$ in absolute value. Is it true that $\bar{\mu}_{\psi_n}$ converge weak-* to the normalized projection of the Haar measure to $Y$?

1.3. **This paper: Quantum unique ergodicity on locally symmetric spaces.** This paper is the first of two papers on this general problem. The main result of the present paper (Theorem 1.6 below) is the construction of the microlocal lift in this setting. We will impose a mild non-degeneracy condition on the sequence of eigenfunctions (see Section 3.3; the assumption essentially amounts to asking that all eigenvalues tend to infinity, at the same rate for operators of the same order.)

With $K$ and $G$ as in Problem 1.5, let $A$ be as in the Iwasawa decomposition $G = NAK$, i.e. $A = \exp(\mathfrak{a})$ where $\mathfrak{a}$ is a maximal abelian subspace of $\mathfrak{p}$. (Full definitions are given in Section 2.1). For $G = \mathrm{SL}_n(\mathbb{R})$ and $K = \mathrm{SO}_n(\mathbb{R})$, one may take $A$ to be the subgroup of diagonal matrices with positive entries. Let $\pi\colon X \to Y$ be the projection. We denote by $dx$ the $G$-invariant probability measures on $X$, and by $dy$ the projection of this measure to $Y$.

The content of the Theorem that follows amounts, roughly, to a "$G$-equivariant microlocal lift" on $Y$.

**Theorem 1.6.** *Let $\{\psi_n\}_{n=1}^{\infty} \subset L^2(Y)$ be a non-degenerate sequence of normalized eigenfunctions, whose eigenvalues approach $\infty$. Then, after replacing $\psi_n$ by an appropriate subsequence, there exist functions $\tilde{\psi}_n \in L^2(X)$ and distributions $\mu_n$ on $X$ such that:*

(1) *The projection of $\mu_n$ to $Y$ coincides with $\bar{\mu}_n$, i.e. $\pi_*\mu_n = \bar{\mu}_n$.*
(2) *Let $\sigma_n$ be the measure $|\tilde{\psi}_n(x)|^2 dx$ on $X$. Then, for every $g \in C_c^{\infty}(X)$, we have $\lim_{n\to\infty}(\sigma_n(g) - \mu_n(g)) = 0$.*
(3) *Every weak-* limit $\sigma_\infty$ of the measures $\sigma_n$ (necessarily a positive measure of mass $\leq 1$) is $A$-invariant.*
(4) *(Equivariance). Let $E \subset \mathrm{End}_G(C^{\infty}(X))$ be a $\mathbb{C}$-subalgebra of bounded endomorphisms of $C^{\infty}(X)$, commuting with the $G$-action. Noting that each $e \in E$ induces an endomorphism of $C^{\infty}(Y)$, suppose that $\psi_n$ is an eigenfunction for $E$ (i.e. $E\psi_n \subset \mathbb{C}\psi_n$). Then we may choose $\tilde{\psi}_n$ so that $\tilde{\psi}_n$ is an eigenfunction for $E$ with the same eigenvalues as $\psi_n$, i.e. for all $e \in E$ there exists $\lambda_e \in \mathbb{C}$ such that $e\psi_n = \lambda_e\psi_n, e\tilde{\psi}_n = \lambda_e\tilde{\psi}_n$.*

We first remark that the distributions $\mu_n$ (resp. the measures $\sigma_n$) generalize the constructions of Zelditch (resp. Wolpert). Although, in view of (2), they carry roughly equivalent information, it is convenient to work with both simultaneously: the distributions $\mu_n$ are canonically defined and easier to manipulate algebraically, whereas the measures $\sigma_n$ are patently positive and are central to the arguments in the sequel to this paper.

*Proof.* For simplicity, we first write the proof in detail for the case where $G$ is simple (the modifications necessary in the general case are discussed in Section 5.1).

In Section 3.2 we define the distributions $\mu_n$. (In the language of Definition 3.3, we take $\mu_n = \mu_{\psi_n}(\varphi_0, \delta)$).

Claim (1) is established in Lemma 3.6.

In Section 3.3 we introduce the non-degeneracy condition. Proposition 3.13 defines $\tilde{\psi}_n$ and establishes the claims (2) and (4). (Observe that this Proposition establishes (2) only for $K$-finite test functions $g$. Since the extension to general $g$ is not necessary for any of our applications, we omit the proof.)

Finally, in section 4 we establish claim (3) (Corollary 4.8) by finding enough differential operators annihilating $\mu_n$. □

*Remark* 1.7.

(1) It is important to verify that non-degenerate sequences of eigenfunctions exist. In the co-compact case (e.g. for the purpose of Theorem 1.10), it was shown in [5, 4] that a positive proportion of the unramified spectrum lies in every open subcone of the Weyl chamber (for definitions see Theorem 2.7 and the discussion in Section 3.1). This is also expected to hold for finite-volume *arithmetic* quotients $Y$. For example, [15, Thm. 5.3] treats the case of $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})/\mathrm{SO}_3(\mathbb{R})$.

(2) We shall use the phrase *non-degenerate quantum limit* to denote any weak-* limit of $\sigma_n$, where notations are as in Theorem 1.6. Note that if $\sigma_\infty$ is such a limit, then claim (2) of the Theorem shows that there exists a subsequence $(n_k)$ of the integers such that $\sigma_\infty(g) = \lim_{n_k\to\infty} \mu_{n_k}(g)$ for all $g \in C_c^\infty(X)$. Depending on the context, we shall therefore use the notation $\sigma_\infty$ or $\mu_\infty$ for a non-degenerate quantum limit.

(3) It is not necessary to pass to a subsequence in Theorem 1.6. See Remark 3.12.

(4) It is likely that the $A$-invariance aspect of Theorem 1.6 could be established by standard microlocal methods; however, the equivariance property does not follow readily from these methods and is

absolutely crucial in applications. It will be used, in the sequel to this paper, in the situation where $E$ is an algebra of endomorphisms generated by Hecke correspondences.

(5) The measures $\sigma_\infty$ all are invariant by the compact group $M = Z_K(\mathfrak{a})$. In fact, Theorem 1.6 should strictly be interpreted as lifting measures to $X/M$ rather than $X$.
(6) Theorem 1.6 admits a natural geometric interpretation. Informally, the bundle $X/M \to Y$ may be regarded as a bundle parameterizing maximal flats in $Y$, and the $A$-action on $X/M$ corresponds to "translation along flats." We refer to Section 5.3 for a further discussion of this point.

The existence of the microlocal lift already places a restriction on the possible weak-* limits of the measures $\{\bar{\mu}_n\}$ on $Y$. In particular, Theorem 1.6 has the following corollary (in this regard see also Remark 1.7(4)).

**Corollary 1.8.** *Let $\{\psi_n\}_{n=1}^\infty \subset L^2(Y)$ be a non-degenerate sequence of normalized eigenfunctions such that $\bar{\mu}_{\psi_n}$ converge in the weak-* topology to a limit measure $\bar{\mu}_\infty$. Then $\bar{\mu}_\infty$ is the projection to $Y$ of an $A$-invariant measure $\mu_\infty$ on $X$. In particular, the support of $\mu_\infty$ must be a union of maximal flats.*

More importantly, Theorem 1.6 allows us to pose a new version of the problem:

**Problem 1.9.** (QUE on homogeneous spaces) In the setting of Problem 1.5, is the $G$-invariant measure on $X$ the unique non-degenerate quantum limit?

1.4. **Arithmetic QUE. Sequel to this paper.** The sequel to this paper will resolve Problem 1.9 for various higher rank symmetric spaces, in the context of *arithmetic* quantum limits. We briefly recall their definition and significance.

Let $Y$ be (for example) a negatively curved manifold. In general, we believe that the multiplicities of the Laplacian $\Delta$ acting on $L^2(Y)$ are quite small, i.e. the $\lambda$-eigenspace has dimension $\ll_\epsilon \lambda^\epsilon$. This question seems extremely difficult even for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, and no better bound is known than the general $O(\lambda^{1/2}/\log(\lambda))$, valid for all negatively curved manifolds.

However, even lacking information on the multiplicities, it transpires that in many natural instances we have a *distinguished basis* for $L^2(Y)$. In that context, it is then natural to ask whether Problem 1.5 or Problem 1.9 can be resolved with respect to this distinguished basis. Since it is believed that the $\Delta$-multiplicities are small, this modification is, philosophically, not too far from the original question. However, it is in many natural cases far more tractable.

The situation of having (something close to) a distinguished basis occurs for $Y = \Gamma\backslash G/K$ and $\Gamma \subset G$ arithmetic. This distinguished basis is obtaining by simultaneously diagonalizing the action of Hecke operators. We shall not give precise definitions here; in any case, we refer to quantum limits arising from subsequences of the distinguished basis as *arithmetic quantum limits*.

In the second paper we apply this results of this paper to the study of arithmetic quantum limits. In particular we settle the conjecture in the case where $\Gamma$ arises from the multiplicative group of a division algebra of prime degree over $\mathbb{Q}$. For brevity, we state the result in the language of automorphic forms; in particular, $\mathbb{A}$ is the ring of adeles of $\mathbb{Q}$.

**Theorem 1.10.** *(QUE for division algebras of prime degree) Let $\mathbb{D}/\mathbb{Q}$ be a division algebra of prime degree $d$ and let $\mathbb{G} = \mathrm{P}\mathbb{D}^\times$ be the associated projective general linear group. Assume that $\mathbb{D}$ is unramified at $\infty$, i.e. that $\mathbb{G}(\mathbb{R}) \simeq \mathrm{PGL}_d(\mathbb{R})$. Let $K_f < \mathbb{G}(\mathbb{A}_f)$ be an open compact subgroup, and let $\Gamma < \mathbb{G}(\mathbb{R})$ be the (congruence) lattice such that $X = \Gamma\backslash\mathbb{G}(\mathbb{R}) \simeq \mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})/K_f$. Then the normalized Haar measure is the unique non-degenerate arithmetic quantum limit on $X$.*

We expect the techniques developed for the proof of Theorem 1.10 will generalize at least to some other locally symmetric spaces, the case of $\mathbb{D}$ being the simplest; but there are considerable obstacles to obtaining a theorem for *any* arithmetic locally symmetric space at present.

Let us make some remarks about the proof of Theorem 1.10. Our approach follows that of Lindenstrauss in [14] which established the above theorem for division algebras of degree 2. This approach is based on result toward the classification of the $A$-invariant measures on $X$. To apply such a result one needs to show further regularity of the limit measure – that $A$ acts on every $A$-ergodic component of $\mu_\infty$ with positive entropy. This was proved for $G = \mathrm{SL}_2$ by Bourgain and Lindenstrauss in [1]. In the higher-rank case we rely on recent results toward the classification of the $A$-invariant measures on $X$, due to Einsiedler-Katok [6], and prove the positive entropy property of $\mu_\infty$.

Establishing positive entropy in higher rank is quite involved. The equivariance (property (4) of Theorem 1.6), applied with $E$ the Hecke algebra, plays a crucial role, just as in [1]. The proof utilizes a study of the behavior of eigenfunctions on Bruhat-Tits buildings and consideration of certain Diophantine questions (these questions are higher-rank versions of the questions: to what extent can CM points of bounded height cluster together?)

1.5. **Acknowledgments.** We are deeply indebted both to P. Sarnak and to E. Lindenstrauss for their encouragement. We have also benefited from discussions with J. Bernstein, A. Reznikov and E. Lapid.

## 2. NOTATION

Section 2.1 defines mostly standard notation and terminology pertaining to semisimple groups and their root systems (we generally follow [12]). Section 2.2 sets up the basic theory of spherical representations; the reader may wish to read at least Definitions 2.3 and 2.6. Section 2.3 defines the various function spaces we will have need of; the notation here is fairly standard.

2.1. **General notation.** Let $G$ denote a non-compact connected simple Lie group with finite center (we discuss generalizations to this in Section 5.1). We choose a Cartan involution $\Theta$ for $G$, and let $K < G$ be the $\Theta$-fixed maximal compact subgroup. Let $S = G/K$ be the symmetric space, with $x_K \in S$ the point with stabilizer $K$. We fix a $G$-invariant metric on $S$. To normalize it, we observe that the tangent space at the point $x_K \in S$ is identified with $\mathfrak{p}$ (see below), and we endow it with the Killing form.

For a lattice $\Gamma < G$ we set $X = \Gamma\backslash G$ and $Y = \Gamma\backslash G/K$, the latter being a locally symmetric space of non-positive curvature. We normalize the Haar measures $dx$ on $X$, $dk$ on $K$ and $dy$ on $Y$ to have total mass 1 (here $dy$ is the pushforward of $dx$ under the the projection from $X$ to $Y$ given by averaging w.r.t. $dk$).

Let $\mathfrak{g} = \mathrm{Lie}(G)$, and let $\theta$ denote the differential of $\Theta$, giving the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ with $\mathfrak{k} = \mathrm{Lie}(K)$. Fix now a maximal abelian subalgebra $\mathfrak{a} \subset \mathfrak{p}$.

We denote by $\mathfrak{a}_{\mathbb{C}}$ the complexification $\mathfrak{a}\otimes_{\mathbb{R}}\mathbb{C}$; we shall occasionally write $\mathfrak{a}_{\mathbb{R}}$ for $\mathfrak{a}$ for emphasis in some contexts. We denote by $\mathfrak{a}^*$ (resp. $\mathfrak{a}^*_{\mathbb{C}}$) the real dual (resp. the complex dual) of $\mathfrak{a}$; again, we shall occasionally write $\mathfrak{a}^*_{\mathbb{R}}$ for $\mathfrak{a}^*$. For $\nu \in \mathfrak{a}^*_{\mathbb{C}}$, we define $\mathrm{Re}(\nu), \mathrm{Im}(\nu) \in \mathfrak{a}^*_{\mathbb{R}}$ to be the real and imaginary parts of $\nu$, respectively.

For $\alpha \in \mathfrak{a}^*$ set $\mathfrak{g}_\alpha = \{X \in \mathfrak{g} | \forall H \in \mathfrak{a} : \mathrm{ad}(H)X = \alpha(H)X \}$, $\Delta(\mathfrak{a} : \mathfrak{g}) = \{\alpha \in \mathfrak{a}^* \setminus \{0\} \mid \mathfrak{g}_\alpha \neq \{0\}\}$ and call the latter the (restricted) *roots* of $\mathfrak{g}$ w.r.t. $\mathfrak{a}$. The subalgebra $\mathfrak{g}_0$ is $\theta$ invariant, and hence $\mathfrak{g}_0 = (\mathfrak{g}_0 \cap \mathfrak{p}) \oplus (\mathfrak{g}_0 \cap \mathfrak{k})$. By the maximality of $\mathfrak{a}$ in $\mathfrak{p}$, we must then have $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathfrak{m}$ where $\mathfrak{m} = Z_{\mathfrak{k}}(\mathfrak{a})$.

The Killing form of $\mathfrak{g}$ induces a standard inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{a}^*$ w.r.t. which $\Delta(\mathfrak{a} : \mathfrak{g}) \subset \mathfrak{a}^*$ is a root system. The associated Weyl group, generated by the root reflections $s_\alpha$, will be denoted $W(\mathfrak{a} : \mathfrak{g})$. This group is also canonically isomorphic to the analytic Weyl groups $N_G(A)/Z_G(A)$ and $N_K(A)/Z_K(A)$. The fixed-point set of any $s_\alpha$ is a hyperplane in $\mathfrak{a}^*$, called a *wall*. The connected components of the complement of the union of the walls are cones, called the *(open) Weyl chambers*. A subset $\Pi \subset \Delta(\mathfrak{a}\colon\mathfrak{g})$ will be called a *system of simple roots* (by abuse of notation a "simple system") if every root can be uniquely expressed as an integral combination of elements of $\Pi$ with either all coefficients non-negative or all coefficients non-positive. For a simple system $\Pi$, the open cone $C_\Pi = \{\nu \in \mathfrak{a}^* \mid \forall\alpha \in \Pi : \langle \nu, \alpha\rangle > 0\}$ is an open Weyl chamber, and the map $\Pi \mapsto C_\Pi$ is a $1-1$ correspondence between simple systems and chambers. The Weyl group acts simply transitively on the chambers and simple systems. The closure of an open chamber will be called a closed chamber. The action of $W(\mathfrak{a}\colon\mathfrak{g})$ on $\mathfrak{a}^*$ extends in the complex-linear way to an action on $\mathfrak{a}^*_\mathbb{C}$ preserving $i\mathfrak{a}^* \subset \mathfrak{a}^*_\mathbb{C}$, and we call an element $\nu \in \mathfrak{a}^*_\mathbb{C}$ *regular* if it is fixed by no $w \in W(\mathfrak{a}\colon\mathfrak{g})$. We use $\rho = \frac{1}{2}\sum_{\alpha>0}(\dim \mathfrak{g}_\alpha)\alpha \in \mathfrak{a}^*$ to denote half the sum of the positive (restricted) roots.

Fixing a simple system $\Pi$ we get a notion of positivity. For $\mathfrak{n} = \oplus_{\alpha>0}\mathfrak{g}_\alpha$ and $\bar{\mathfrak{n}} = \Theta\mathfrak{n}$ we have $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{m} \oplus \bar{\mathfrak{n}}$ and (Iwasawa decomposition) $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{k}$. By means of the Iwasawa decomposition, we may uniquely write every $X \in \mathfrak{g}$ in the form $X = X_\mathfrak{n} + X_\mathfrak{a} + X_\mathfrak{k}$. We sometimes also write $H_0(X)$ for $X_\mathfrak{a}$.

Let $N, A < G$ be the subgroups corresponding to the subalgebras $\mathfrak{n}, \mathfrak{a} \subset \mathfrak{g}$ respectively, and let $M = Z_K(\mathfrak{a})$. Then $A$ is a maximal split torus in $G$, and $\mathfrak{m} = \mathrm{Lie}(M)$, though $M$ is not necessarily connected. Moreover $P_0 = NAM$ is a minimal parabolic subgroup of $G$, with the map $N \times A \times M \to P_0$ being a diffeomorphism. The map $N \times A \times K \to G$ is a (surjective) diffeomorphism (Iwasawa decomposition), so for $g \in G$ there exists a unique $H_0(g) \in \mathfrak{a}$ such that $g = n\exp(H_0(g))k$ for some $n \in N$, $k \in K$. The map $H_0 : G \to \mathfrak{a}$ is continuous; restricted to $A$ it is the inverse of the exponential map.

Let $\mathfrak{g}_\mathbb{C} = \mathfrak{g} \otimes_\mathbb{R} \mathbb{C}$ denote the complexification of $\mathfrak{g}$. It is a complex semi-simple Lie algebra. Let $\theta_\mathbb{C}$ denote the *complex-linear* extension of $\theta$ to $\mathfrak{g}_\mathbb{C}$. It is *not* a Cartan involution of $\mathfrak{g}_\mathbb{C}$. We fix a maximal abelian subalgebra $\mathfrak{b} \subset \mathfrak{m}$ and set $\mathfrak{h} = \mathfrak{a} \oplus \mathfrak{b}$. Then $\mathfrak{h}_\mathbb{C} = \mathfrak{h} \otimes \mathbb{C} \subset \mathfrak{g}_\mathbb{C}$ is a Cartan subalgebra, with the associated root system $\Delta(\mathfrak{h}_\mathbb{C}\colon\mathfrak{g}_\mathbb{C})$ satisfying $\Delta(\mathfrak{a}\colon\mathfrak{g}) = \{\alpha\restriction_\mathfrak{a}\}_{\alpha\in\Delta(\mathfrak{h}_\mathbb{C}\colon\mathfrak{g}_\mathbb{C})} \setminus \{0\}$. Moreover, we can find a system of simple roots $\Pi_\mathbb{C} \subset \Delta(\mathfrak{h}_\mathbb{C}\colon\mathfrak{g}_\mathbb{C})$ and a system of simple roots $\Pi \subset \Delta(\mathfrak{a}\colon\mathfrak{g})$ such that the positive roots w.r.t. $\Pi$ are precisely the nonzero restrictions of the positive roots w.r.t. $\Pi_\mathbb{C}$. We fix such a compatible pair of simple systems, and let $\rho_\mathfrak{h}$ denote half the sum of the roots in $\Delta(\mathfrak{h}_\mathbb{C}\colon\mathfrak{g}_\mathbb{C})$, positive w.r.t. $\Pi_\mathbb{C}$.

Let $F_0 \subset \Delta(\mathfrak{h}_\mathbb{C}\colon\mathfrak{g}_\mathbb{C})$ consist of the roots that restrict to $0$ on $\mathfrak{a}$, $F_0^+ \subset F_0$ those positive w.r.t. $\Pi_\mathbb{C}$. Let $\mathfrak{n}_M = \oplus_{\alpha\in F_0^+}(\mathfrak{g}_\mathbb{C})_\alpha$, $\bar{\mathfrak{n}}_M = \oplus_{\alpha\in F_0^+}(\mathfrak{g}_\mathbb{C})_{-\alpha}$. Then $\mathfrak{m}_\mathbb{C} = \mathfrak{n}_M \oplus \mathfrak{b}_\mathbb{C} \oplus \bar{\mathfrak{n}}_M$ and $\mathfrak{g}_\mathbb{C} = \mathfrak{n}_\mathbb{C} \oplus \mathfrak{n}_M \oplus \mathfrak{h}_\mathbb{C} \oplus \bar{\mathfrak{n}}_M \oplus \bar{\mathfrak{n}}_\mathbb{C}$.

For $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, set $\|\nu\|^2 = \langle \mathrm{Re}(\nu), \mathrm{Re}(\nu)\rangle + \langle \mathrm{Im}(\nu), \mathrm{Im}(\nu)\rangle$ (with the inner products taken in $\mathfrak{a}_{\mathbb{R}}^*$).

If $\mathfrak{l}_{\mathbb{C}}$ is a complex Lie algebra, then we denote by $U(\mathfrak{l}_{\mathbb{C}})$ its universal enveloping algebra, and by $\mathfrak{Z}(\mathfrak{l}_{\mathbb{C}})$ its center. In particular we set $\mathfrak{Z} = \mathfrak{Z}(\mathfrak{g}_{\mathbb{C}})$.

2.2. **Spherical Representations and the model $(V_K, I_\nu)$.** We recall some facts from the representation theory of compact and semi-simple groups. At the end of this section we analyze a model (the "compact picture") for the spherical dual of $G$.

**Theorem 2.1.** [12, Th. 1.12] *Let $K$ be a compact topological group and let $\hat{K}_{\mathrm{fin}}$ be the set of equivalence classes of irreducible finite-dimensional unitary representations of $K$.*

(1) *(Peter-Weyl) Every $\rho \in \hat{K}_{\mathrm{fin}}$ occurs discretely in $L^2(K)$ with multiplicity equal to its dimension $\mathrm{d}(\rho)$. Moreover, $L^2(K)$ is isomorphic to the Hilbert direct sum of its isotypical components $\{L^2(K)_\rho\}_{\rho\in\hat{K}_{\mathrm{fin}}}$.*
(2) *Let $\pi : K \to \mathrm{GL}(W)$ be a representation of $K$ on the locally convex complete space $W$. Then $\oplus_{\rho\in\hat{K}} W_\rho$ is dense in $W$, where $W_\rho$ is the $\rho$-isotypical subspace.*
(3) *Every irreducible representation of $K$ on a locally convex, complete space is finite-dimensional and hence unitarizable. In particular, $\hat{K}_{\mathrm{fin}}$ is the unitary dual of $K$.*
(4) *For $K$ as in Section 2.1, $\hat{K}$ is countable.*

Note that while [12, Th. 1.12(c-e)] are only claimed for unitary representations on Hilbert spaces, their proofs only rely on the action of the convolution algebra $C(K)$ on representations of $K$, and hence carry over with little modification to the more general context needed here. The last conclusion follows from the separability of $L^2(K)$, which in turn follows from the separability of $K$.

*Notation* 2.2. Let $\pi : K \to \mathrm{GL}(W)$ be as above. The algebraic direct sum $W_K \stackrel{\mathrm{def}}{=} \oplus_{\rho\in\hat{K}} W_\rho$ consists precisely of these $w \in W$ which generate a finite-dimensional $K$-subrepresentation. We refer to $W_K$ as the space of $K$-finite vectors. We will use $W^K$ to denote these vectors of $W$ fixed by $K$.

**Definition 2.3.** *Set $V = L^2(M\backslash K)$, and set $V_K \subset V$ to be the space of $K$-finite vectors. Let $C^\infty(M\backslash K)$ be the smooth subspace, $C^\infty(M\backslash K)'$ the space of distributions on $M\backslash K$. Let $V_K'$ (resp. $V'$) be the dual to $V_K$ (resp. $V$). Then we have natural inclusions $V_K \subset C^\infty(M\backslash K) \subset V$ and $V_K' \supset C^\infty(M\backslash K)' \supset V'$; further, we have (Riesz representation) a conjugate-linear isomorphism*

$$V \stackrel{T}{\hookrightarrow} V' \tag{2.1}$$

*where the map $T\colon V \to V'$ is defined via the rule $T(f)(g) = \langle g, f\rangle_V = \int_{M\backslash K} g\overline{f}dk$.*

*Fix an increasing exhaustive sequence of finite dimensional $K$-stable subspaces of $V_K$, i.e. a sequence $V_1 \subset V_2 \subset \cdots \subset V_N \subset V_{N+1} \subset \ldots$ of subspaces such that $\cup_{i=1}^{\infty} V_i = V_K$ and each $V_i$ is a $K$-subrepresentation.*

*For $\Phi \in V_K'$ and $1 \leq N \in \mathbb{Z}$, define the $N$-truncation of $\Phi$ as the unique element $\Phi_N \in V_N$ such that $T(\Phi_N) - \Phi$ annihilates $V_N$.*

*Finally let $\varphi_0 \in V_K$ be the function that is identically* $1$.

**Definition 2.4.** *Let $\mu$ be a regular Borel measure on a space $X$. Call a sequence of non-negative functions $\{f_j\} \in L^1(\mu)$ a* $\delta$-sequence *at $x \in X$ if, for every $j$, $\int f_j d\mu = 1$, and moreover if, for every $g \in C(X)$, $\lim_{j\to\infty} \int f_j \cdot g d\mu = g(x)$.*

**Lemma 2.5.** *There exists a sequence $\{f_j\}_{j=1}^{\infty} \subset V_K$ such that $|f_j|^2$ is a $\delta$-sequence on $M\backslash K$.*

*Proof.* Let $\{h_j\}_{j=1}^{\infty} \subset C(M\backslash K)$ be a $\delta$-sequence. By the Peter-Weyl theorem $V_K$ is dense in $C(M\backslash K)$, so that for every $j$ we can choose $f_j' \in V_K$ such that $\left\| \sqrt{h_j(k)} - f_j'(k) \right\|_{\infty} \leq \frac{1}{2^j}$. Then one may take $f_j = \frac{f_j'}{\left\|f_j'\right\|_2}$. □

Secondly, we recall the construction of the spherical principal series representations of a semi-simple Lie group. An irreducible representation of $G$ is *spherical* if it contains a $K$-fixed vector. Such a vector is necessarily unique up to scaling.

To any $\nu \in \mathfrak{a}_{\mathbb{C}}^*$ we associate the character $\chi_\nu(p) = \exp(\nu(H_0(p))$ of $P_0$ and the induced representation with $(\mathfrak{g}, K)$-module

(2.2)
$$\operatorname{Ind}_{P_0}^G \chi_\nu = \left\{ f \in C^\infty(G)_K \mid \forall p \in P, g \in G : f(pg) = e^{\langle \nu+\rho, H_0(p)\rangle} f(g) \right\}$$

By the Iwasawa decomposition, every $f \in \operatorname{Ind}_{P_0}^G \chi_\nu$ is determined by its restriction to $K$; this restriction defines an element of the space $V_K$. Conversely, every $f \in V_K$ extends uniquely to a member of $\operatorname{Ind}_{P_0}^G \chi_\nu$.

**Definition 2.6.** *For $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, we denote by $(I_\nu, V_K)$ the representation of $\mathfrak{g}$ on $V_K$ fixed by the discussion above; we shall also use $I_\nu$ to denote the corresponding action of $\mathfrak{g}$ on $C^\infty(M\backslash K)$ and of $G$ on $V$. We shall denote by $I_\nu'$ the dual action of $\mathfrak{g}$ on either $V_K'$ or $C^\infty(M\backslash K)'$.*

Note also that $\varphi_0 \in V_K$ (see Definition 2.3) is a spherical vector for the representation $(I_\nu, V_K)$.

**Theorem 2.7.** *(The unitary spherical dual; references are drawn from* [12]*)*

(1) *For any $\nu \in \mathfrak{a}_{\mathbb{C}}^*$, $\operatorname{Ind}_{P_0}^G \chi_\nu$ has a unique spherical irreducible subquotient, to be denoted $\pi_\nu$. [Th. 8.37] Any spherical irreducible*

*unitary representation of $G$ is isomorphic to $\pi_\nu$ for some $\nu$. [Th. 8.38] We have $\pi_{\nu_1} \simeq \pi_{\nu_2}$ iff there exists $w \in W(\mathfrak{a}\colon\mathfrak{g})$ such that $\nu_2 = w\nu_1$.*

(2) *[§7.1-3] If $\mathrm{Re}(\nu) = 0$ then $\mathrm{Ind}_{P_0}^G \chi_\nu$ is unitarizable, with the invariant Hermitian form given by $\langle f, g\rangle = \int_{M\backslash K} f(k)\overline{g(k)}dk$. This representation has a unique spherical summand (necessarily isomorphic to $\pi_\nu$), and we let $j_\nu\colon V_K \to \pi_\nu$ denote the orthogonal projection map. [Th. 7.2] If $\nu$ is regular then $\mathrm{Ind}_{P_0}^G \chi_\nu$ is irreducible.*

(3) *[§16.5(7) & Th. 16.6] If $\pi_\nu$ is unitarizable then $\mathrm{Re}(\nu)$ belongs to the convex hull of $\{w\rho\}_{\{w\in W(\mathfrak{a}\colon\mathfrak{g})\}} \subset \mathfrak{a}^*$, a compact set. Moreover, there exists $w \in W(\mathfrak{a}\colon\mathfrak{g})$ such that $w^2 = 1$ and $w\nu = -\bar{\nu}$. In particular if $\mathrm{Re}(\nu) \neq 0$, then $w \neq 1$, and since Im($\nu$) is $w$-fixed it is not regular.*

Note that the norm on $\pi_\nu$ is only unique up to scaling. If $\mathrm{Re}(\nu) = 0$ and $\mathrm{Im}(\nu)$ is regular (the main case under consideration), we choose $\|\varphi_0\|_{\pi_\nu} = 1$.

For future reference we compute the action of $\mathfrak{g}$ on $V_K$ via $I_\nu$. First, remark that the action of $K$ on $V = L^2(M\backslash K)$ is given by right translation, and the action of $\mathfrak{k} \subset \mathfrak{g}$ on $V_K$ is then given by right differentiation.

Secondly, recall that if $U \subset \mathbb{R}^n$ is open, a *differential operator* $\mathbf{D}$ on $U$ is an expression of the form $\sum_{i=1}^K f_i\partial_1^{\alpha_1}\dots\partial_n^{\alpha_n}$, where the $f_i$ are smooth and $\alpha_j \geq 0$. If $M$ is a smooth $n$-manifold, we say a map $\mathbf{D} : C^\infty(M) \to C^\infty(M)$ is a differential operator if it is defined by a differential operator in each coordinate chart.

**Lemma 2.8.** *Let $f \in V_K$ and let $X \in \mathfrak{g}$. Then there exists a differential operator $\mathbf{D}_X$ on $M\backslash K$ (depending linearly on $X$ and independent of $\nu$) such that for every $k \in K$,*

$$(I_\nu(X)f)(k) = \langle \nu + \rho, H_0(\mathrm{Ad}(k)X)\rangle f(k) + (\mathbf{D}_X f)(k).$$

*Proof.* Let $t \in \mathbb{R}$ be small, and consider $f(k\exp(tX)) = f(\exp(t\,\mathrm{Ad}(k)X)\cdot k)$. We write the Iwasawa decomposition of $\mathrm{Ad}(k)X \in \mathfrak{g}$ as $\mathrm{Ad}(k)X = X_\mathfrak{n}(k) + X_\mathfrak{a}(k) + X_\mathfrak{k}(k)$ where $X_\mathfrak{a}(k) = H_0(\mathrm{Ad}(k)X)$. By the Baker-Campbell-Hausdorff formula, $\exp(t\,\mathrm{Ad}(k)X) = \exp(tX_\mathfrak{n}(k))\cdot\exp(tX_\mathfrak{a}(k))\cdot \exp(tX_\mathfrak{k}(k)) + O(t^2)$, so that:

$$(I_\nu(X)f)(k) = \frac{d}{dt}f\left(\exp(tX_\mathfrak{n}(k))\cdot\exp(tX_\mathfrak{a}(k))\cdot k\right)\restriction_{t=0} + \frac{d}{dt}f\left(\exp(tX_\mathfrak{k}(k))k\right)\restriction_{t=0}.$$

To conclude, observe that $f \mapsto \frac{d}{dt}f\left(\exp(tX_\mathfrak{k}(k))k\right)\restriction_{t=0}$ defines a differential operator $\mathbf{D}_X$ on $M\backslash K$. □

Lemma 2.8 will be used in the following way: as $\|\nu\| \to \infty$, the operator $I_\nu(\frac{X}{\|\nu\|})$ acts on $V_K$ in a very simple fashion, *modulo* certain error terms of order $\|\nu\|^{-1}$. The simplicity of this "rescaled" action as $\|\nu\| \to \infty$ will be of importance in our analysis.

2.3. **Some functional analysis.** We collect here some simple functional analysis facts that we shall have need of.

Let $C_c^\infty(X)$ denote the space of smooth functions of compact support on $X$. It is endowed with the usual "direct-limit" topology: fix a sequence of $K$-invariant compact sets $C_1 \subset C_2 \subset \ldots$ such that their interiors exhaust $X$. Then the $C_c^\infty(C_i)$ exhaust $C_c^\infty(X)$. $C_c^\infty(C_i)$ is endowed as usual with a family of seminorms, viz. for any $\mathcal{D} \in U(\mathfrak{g}_{\mathbb{C}})$ we define $\|f\|_{C_i,\mathcal{D}} = \sup_{x\in C_i} |\mathcal{D}f|$. These seminorms induce a topology on each $C_c^\infty(C_i)$. We give $C_c^\infty(X)$ the topology of the union of $C_c^\infty(C_i)$, i.e. a map from $C_c^\infty(X)$ is continuous if and only if its restriction to each $C_c^\infty(C_i)$ is continuous.

In other words: a sequence of functions converges in $C_c^\infty(X)$ if their supports are all contained in a fixed compact set, and all their derivatives converge uniformly on that compact set.

$C_c^\infty(X)$ is a locally convex complete space in this topology. In particular, its subspace $C_c^\infty(X)_K$ of $K$-finite vectors is dense. We denote by $C_c^\infty(X)'$ (resp. $C_c^\infty(X)'_K$) the topological dual to $C_c^\infty(X)$ (resp. the algebraic dual to $C_c^\infty(X)_K$). Both spaces will be endowed with the weak-* topology. We shall refer to an element of $C_c^\infty(X)'$ as a *distribution* on $X$.

Let $C_0(X)$ be the Banach space of continuous functions on $X$ decaying at infinity, endowed with the supremum norm. Let $C_0(X)'$ be the continuous dual of $C_0(X)$; the Riesz representation theorem identifies it with the space of finite (signed) Borel measures on $X$. We endow $C_0(X)'$ with the weak-* topology.

It is easy to see that $C_c^\infty(X)_K$ is dense in $C_0(X)$. In particular any (algebraic) linear functional on $C_c^\infty(X)_K$ which is bounded w.r.t. the $\sup$-norm extends to a finite signed measure on $X$, with total variation equal to the norm of the functional. Moreover, if this functional is non-negative on the non-negative members of $C_c^\infty(X)_K$ then the associated measure is a positive measure.

## 3. Representation-Theoretic Lift

3.1. **Introduction and motivation.** Suppose $\psi \in L^2(Y)$ has $\|\psi\|_2 = 1$ and an eigenfunction of $\mathfrak{Z}$. The aim of the present section is to construct a distribution $\mu_\psi$ on $Y$ that lifts the measure $\bar{\mu}_\psi$ on $Y$, and establish some basic properties of $\mu_\psi$.

In the situation of Theorem 1.6, if $\psi = \psi_n$, the corresponding distribution will be the distributions $\mu_n$ discussed in the proof of Theorem 1.6. The functions $\tilde{\psi}_n$ will then be chosen so that the measures $|\tilde{\psi}_n(x)|^2 dx$ approximate $\mu_n$; finally, both $|\tilde{\psi}_n(x)|^2$ and $\mu_n$ will become $A$-invariant as $n \to \infty$.

We begin by fixing notation and providing some motivation for the relatively formal definitions that follow.

Setting $\psi(x) = \psi(xK)$ for any $x \in X$, we can think of $\psi$ as a function on $X$. By the uniqueness of spherical functions [7, Th. 4.3 & 4.5], $\psi$ generates an irreducible spherical $G$-subrepresentation of $L^2(X)$. As discussed in Section 2.2, we can then find $\nu \in \mathfrak{a}_{\mathbb{C}}^*$ such that this representation is isomorphic to $\pi_\nu$ (in particular, $\pi_\nu$ is unitarizable). We will assume for the rest of this section that $Re(\nu) = 0$, i.e. $\pi_\nu$ is *tempered*, and that $\nu$ is regular. This will eventually be the only case of interest to us in view of the non-degeneracy assumption made later (Definition 3.8). In this case $(V_K, I_\nu)$ is irreducible and isomorphic to $\pi_\nu$. It follows that there is a unique $G$-homomorphism $R_\psi : (V, I_\nu) \to L^2(X)$ such that $R_\psi(\varphi_0) = \psi$. The normalization $\|\psi\|_{L^2(X)} = 1$ now implies $\|R_\psi(f)\|_{L^2(X)} = \|f\|_{L^2(K)}$ for any $f \in V_K$, i.e. that $R_\psi$ is an isometry.

We now give the rough idea of the construction that follows in the language of Wolpert and Lindenstrauss; the language we shall use later is slightly different, so the discussion here also provides a translation. The strategy of proof is similar to theirs; in a sense, the main difficulty is finding the "correct" definitions in higher rank. For instance, the proofs of Wolpert and Lindenstrauss use heavily the fact that $K$-types for $\mathrm{PSL}_2(\mathbb{R})$ have multiplicity one, and the explicit action of the Lie algebra by raising and lowering operators. We shall need a more intrinsic approach to handle the general case.

The measure $\bar{\mu}_\psi$ on $Y$ is defined by $g \mapsto \int_X g(x)|\psi(x)|^2 dx$. More generally, suppose that $\psi' \in L^2(X)$ belongs to the $G$-subrepresentation generated by $\psi$, i.e. $\psi' \in R_\psi(V)$. We can then consider the (signed) measure

$$\sigma : g \mapsto \int_X \psi(x)\overline{\psi'(x)}g(x)dx. \tag{3.1}$$

If $g(x)$ is $K$-invariant, then so is the product $\psi(x)g(x)$, and it follows that the right-hand side of (3.1) depends only on the projection of $\psi'$ onto $R_\psi(V)^K$. The space $R_\psi(V)^K$ is one-dimensional, spanned by $\psi$, and it follows that if $\psi' - \psi \perp \psi$ then the measure $\sigma$ on $X$ projects to the measure $\bar{\mu}_\psi$ on $Y$.

The distribution $\mu_\psi$ we shall be construct will be in the spirit of (3.1), but with $\psi'$ a "generalized vector" in $R_\psi(V)$. Suppose, in fact, that $\psi'_1, \psi'_2, \dots \psi'_n, \dots$ are an infinite sequence of elements of $R_\psi(V)$ that transform under different $K$-types, and suppose further that $g \in C_c^\infty(X)_K$. Then, by considering $K$-types, the integral $\int_X \psi(x)\overline{\psi'_j(x)}g(x)dx$ vanishes for all sufficiently large $j$. It follows that, if one sets $\psi'$ to be the *formal sum* $\sum_{j=1}^\infty \psi'_j$, one can make sense of (3.1) by interpreting it as:

$$\sigma(g) = \sum_{j=1}^{\infty} \int_X \psi(x)\overline{\psi_j'(x)}g(x)dx$$

In other words, if $g \in C_c^\infty(X)_K$, we may make sense of (3.1) while allowing $\psi'$ to belong to the space $\widehat{V_K}$ of "infinite formal sums of $K$-types." Our definition of $\mu_\psi$ will, indeed, be of the form (3.1) but with $\psi'$ an "infinite formal sum" of this kind.

For a certain choice of $\psi'$ (denoted $\Phi_\infty$ in [13]), we will wish to show that (3.1) is "approximately a positive measure" and "approximately $A$-invariant," where both statements become true in the large eigenvalue limit in an appropriate sense. For the "approximate positivity," we shall integrate (3.1) by parts to show that there exists another unit vector $\psi'' \in R_\psi(V)$ such that $\int_X \psi(x)\overline{\psi'(x)}g(x)dx \approx \int_X |\psi''(x)|^2 g(x)dx$, where the right-hand side is evidently a positive measure. For the "approximate $A$-invariance," we will construct differential operators that annihilate $\psi(x)\overline{\psi'(x)}$; this reduces to a purely algebraic question of constructing elements in $U(\mathfrak{g})$ that annihilate a vector in a certain tensor product representation.

The space $\widehat{V_K}$ is very closely linked to the dual $V_K'$ of the $K$-finite vectors: the conjugate linear isomorphism $T : V \to V'$ (2.1) extends to a conjugate-linear isomorphism $T : \widehat{V_K} \to V_K'$. For formal reasons, it is simpler to work with $V_K'$ than $\widehat{V_K}$; this is the viewpoint we shall take in Definition 3.1. To motivate this viewpoint, let us rewrite (3.1) in a different fashion. Let $v' \in V$ be chosen so that $\psi' = R_\psi(v')$, and let $P$ be the orthogonal projection of $L^2(X)$ onto $R_\psi(V)$. We may rewrite (3.1) – using the notations of Definition 2.3 – as follows:

$$\begin{aligned}\sigma(g) &= \langle \psi(x)g(x), \psi'(x)\rangle_{L^2(X)} = \langle P(\psi(x)g(x)), \psi'(x)\rangle_{L^2(X)} \\ &= \langle R_\psi^{-1} \circ P(\psi(x)g(x)), v'\rangle_V = T(v') \circ R_\psi^{-1} \circ P(\psi(x)g(x))\end{aligned} \tag{3.2}$$

Now, if $g \in C_c^\infty(X)_K$, then the quantity $R_\psi \circ P(\psi(x)g(x))$ is $K$-finite, i.e. belongs to $V_K$. It follows that, if $g \in C_c^\infty(X)_K$, the last expression of (3.2) makes formal sense if we replace $T(v')$ by any functional $\Phi \in V_K'$.

### 3.2. Lifting a single (non-degenerate) eigenfunction.

**Definition 3.1.** *Let $\Phi \in V_K'$ be an (algebraic) functional, and $f \in V_K$. Let $\mu_\psi(f, \Phi)$ be the functional on $C_c^\infty(X)_K$ defined by the rule:*

$$\mu_\psi(f, \Phi)(g) = \Phi \circ R_\psi^{-1} \circ P(R_\psi(f) \cdot g) \tag{3.3}$$

*where $g \in C_c^\infty(X)_K$, $P : L^2(X) \to R_\psi(V)$ is the orthogonal projection, and $R_\psi(f) \cdot g$ denotes pointwise multiplication of functions on $X$.*

*Remark* 3.2. In fact, if $\Phi \in C^\infty(M\backslash K)'$ (see equation (2.1)) then $\mu_\psi(f, \Phi)$ extends to an element of $C_c^\infty(X)'$, i.e. defines a distribution on $X$: $\mu_\psi$ is the composite

$$C_c^\infty(X) \xrightarrow{g \mapsto R_\psi(f)g} C_c^\infty(X) \xrightarrow{R_\psi^{-1}P} C^\infty(M\backslash K) \xrightarrow{\Phi} \mathbb{C},$$

and it is easy to verify that each of these maps is continuous. This is never used in our arguments: we use this observation only to refer to certain $\mu_\psi$ as "distributions".

**Definition 3.3.** *Let $\delta \in V_K'$ be the distribution $\delta(f) = f(1)$, and call $\mu_\psi \stackrel{\text{def}}{=} \mu_\psi(\varphi_0, \delta)$ the (non-degenerate)* microlocal lift *of $\bar{\mu}_\psi$.*

The rest of the section will exhibit basic formal properties of this definition. We will establish most of the formal properties of $\mu_\psi$ by restricting $\Phi$ to be of the form $T(f_2)$, where the conjugate-linear mapping $T$ is as defined in (2.1). This situation will occur sufficiently often that, for typographical ease, it will be worth making the following definition:

**Definition 3.4.** *Let $f_1, f_2 \in V_K$. We then set $\mu_\psi^T(f_1, f_2) = \mu_\psi(f_1, T(f_2))$.*

**Lemma 3.5.** *Suppose $f_1, f_2 \in V_K$. Then*

$$\mu_\psi^T(f_1, f_2)(g) = \int_X R_\psi(f_1)(x)\overline{R_\psi(f_2)(x)}g(x)dx. \tag{3.4}$$

*and $\mu_\psi^T$ defines a signed measure on $X$ of total variation at most $||f_1||_{L^2(K)}||f_2||_{L^2(K)}$. If $f_1 = f_2$, then $\mu_\psi^T(f_1, f_1)$ is a positive measure of mass $||f_1||_{L^2(K)}^2$.*

*Proof.* (3.4) is a consequence of the definition of $\mu$. The Cauchy-Schwarz inequality implies that $|\mu_\psi^T(f_1, f_2)(g)| \leq ||f_1||_{L^2(K)}||f_2||_{L^2(K)}||g||_{L^\infty(X)}$, whence the second conclusion. The last assertion is immediate. □

In fact, it may be helpful to think of $\mu_\psi$ as being given by a distributional extension of the formula (3.4); see the discussion of Section 3.1.

**Lemma 3.6.** *The distribution $\mu_\psi(\varphi_0, \delta)$ on $X$ projects to the measure $|\psi|^2 dy$ on $Y$.*

*Proof.* In view of the previous Lemma, it will suffice to show that the distribution $\mu_\psi(\varphi_0, \delta) - \mu_\psi^T(\varphi_0, \varphi_0)$ on $X$ projects to $0$ on $Y$. This amounts to showing that $\mu_\psi(\varphi_0, \delta - T(\varphi_0))$ annihilates any $K$-invariant function $g \in C_c^\infty(X)^K$. Taking into account that the functional $\delta - T(\varphi_0)$ on $V_K$ annihilates any $K$-invariant vector, the claim follows from the definition of $\mu_\psi$. □

**Lemma 3.7.** *The map $\mu_\psi : V_K \otimes V_K' \to C_c^\infty(X)_K'$ is equivariant for the natural $\mathfrak{g}$-actions on both sides.*

*Proof.* This follows directly from the definition of $\mu_\psi$. $\square$

Concretely speaking, this says that for $f \in V_K, \Phi \in V_K', g \in C_c^\infty(X)_K, X \in \mathfrak{g}$ we have

$$\mu_\psi(Xf_1, \Phi)(g) + \mu_\psi(f_1, X\Phi)(g) + \mu_\psi(f_1, \Phi)(Xg) = 0 \tag{3.5}$$

where $X$ acts on $V_K$ via $I_\nu$ and on $V_K'$ via $I_\nu'$. In particular, if $f_1, f_2 \in V_K$ we have

$$\mu_\psi^T(Xf_1, f_2)(g) + \mu_\psi^T(f_1, Xf_2)(g) + \mu_\psi^T(f_1, f_2)(Xg) = 0 \tag{3.6}$$

3.3. **Sequences of eigenfunctions and quantum limits.** In what follows we shall consider $\{\psi_n\}_{n=1}^\infty \subset L^2(Y)$, a sequence of eigenfunctions with parameters $\{\nu_n\}$ diverging to $\infty$ (i.e. leaving any compact set). Set $\tilde{\nu}_n = \frac{\nu_n}{||\nu_n||}$. For $f_1, f_2 \in V_K$ and $\Phi \in V_K'$, we abbreviate $\mu_{\psi_n}^T(f_1, f_2)$ (resp. $\mu_{\psi_n}(f, \Phi)$) to $\mu_n^T(f_1, f_2)$ (resp. $\mu_n(f, \Phi)$), and we abbreviate the microlocal lift $\mu_{\psi_n}$ ($:= \mu_n(\varphi_0, \delta)$) to $\mu_n$.

**Definition 3.8.** *($G$ simple) We say a sequence $\psi_n$ is* non-degenerate *if every limit point of the sequence $\tilde{\nu}_n$ is regular.*

*We say that it is* conveniently arranged *if it is nondegenerate, $\lim_{n\to\infty} \tilde{\nu}_n$ exists, $Re(\nu_n) = 0$ for all $n$, the $\nu_n$ are all regular, and for all $f_1, f_2 \in V_K$ the measures $\mu_n^T(f_1, f_2)$ converge in $C_0(X)'$ as $n \to \infty$. In this situation we denote $\lim_{n\to\infty} \tilde{\nu}_n$ by $\tilde{\nu}_\infty$.*

The existence of non-degenerate sequences of eigenfunctions was discussed in Remark 1.7. This follows from strong versions of Weyl's Law on $Y$. By Theorem 2.7, the non-degeneracy of a sequence $\psi_n$ as in the Definition implies $\mathrm{Re}(\nu_n) = 0$ for all large enough $n$. For fixed $f_1, f_2 \in V_K$ the total variation of the measures $\mu_n^T(f_1, f_2)$ is bounded independently of $n$ (Lemma 3.5); in view of the (weak-*) compactness of the unit ball in $C_0(X)'$ it follows that this sequence of measures has a convergent subsequence. Combining this remark with the fact that $V_K$ has a countable basis, a diagonal argument shows that every non-degenerate sequence of eigenfunctions has a conveniently arranged subsequence.

Now suppose $\{\psi_n\}$ is a conveniently arranged sequence and fix $f_1 \in V_K, \Phi \in V_K', g \in C_c^\infty(X)_K$. Let $\Phi_N$ be the $N$-truncation of $\Phi$ (see Definition 2.3). In view of (3.3), if we choose $N := N(f_1, g)$ sufficiently large, then $\mu_n(f_1, \Phi)(g) = \mu_n^T(f_1, \Phi_N)(g)$. It follows that the limit $\lim_{n\to\infty} \mu_n(f_1, \Phi)(g)$ exists.

We may consequently define $\mu_\infty \colon V_K \times V_K' \to C_c^\infty(X)_K'$ and $\mu_\infty^T \colon V_K \times V_K \to C_c^\infty(X)_K'$ by the rules:

$$\begin{aligned} \mu_\infty(f, \Phi)(g) &= \lim_{n\to\infty} \mu_n(f_1, \Phi)(g), \quad (g \in C_c^\infty(X)_K) \\ \mu_\infty^T(f_1, f_2) &= \mu_\infty(f_1, T(f_2)) \end{aligned} \tag{3.7}$$

**Lemma 3.9.** *For fixed $f_1 \in V_K$, the map $\Phi \to \mu_\infty(f_1, \Phi)$ is continuous as a map $V_K' \to C_c^\infty(X)_K'$, both spaces being endowed with the weak topology.*

*Proof.* This is an easy consequence of the definitions. □

It is natural to ask whether $\mu_\infty(f_1, \Phi)$ extends to an element of $C_c^\infty(X)'$, especially if $\Phi \in C^\infty(M\backslash K)'$. Indeed, it is possible to make quantitative the argument of Remark 3.2 to obtain a uniform bound on the distributions $\mu_n(f_1, \Phi)$. This will not be needed in this paper, however, since for our choice of $(f_1, \Phi)$, the limiting distribution is positive (in particular a measure), a fact we will prove directly.

Henceforth $\{\psi_n\}_{n=1}^\infty$ will be a conveniently arranged sequence. We will show that $\mu_\infty(\varphi_0, \delta)$ is positive and bounded w.r.t. the $L^\infty$ norm on $C_c^\infty(X)_K$. It hence extends to a finite positive measure.

*Remark.* In the case of a semisimple group, one can allow the projection of the parameter to each simple factor to tend to infinity at a different rate. The definition of a non-degenerate limit can then remain unchanged. The $\tilde{\nu}_n$ however must be defined with greater care – see Section 5.1.

The key to the positivity of the limits is the following lemma (cf. [21, Prop. 3.3], [13, Th. 3.1]).

**Lemma 3.10.** *(Integration by parts) Let $\{\psi_n\}$ be conveniently arranged. Then, for any $f,\ f_1,\ f_2 \in V_K$ we have:*

$$\mu_\infty^T(f_1, f \cdot f_2) = \mu_\infty^T(\bar{f} \cdot f_1, f_2). \tag{3.8}$$

*Here e.g. $f \cdot f_2$ denotes pointwise multiplication of functions on $M\backslash K$.*

*Proof.* We start by exhibiting explicit functions $f$ for which (3.8) is valid.

Extend every $\nu \in \mathfrak{a}_\mathbb{C}^*$ to $\mathfrak{g}_\mathbb{C}^*$ via the Iwasawa decomposition $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{a} \oplus \mathfrak{k}$. For any $X \in \mathfrak{g}$, let $p_X(k) = \frac{1}{i}\langle \tilde{\nu}_\infty, \mathrm{Ad}(k)X\rangle$. For fixed $X$, $k \mapsto p_X(k)$ defines a $K$-finite element of $L^2(M\backslash K)$.

By (3.6), for every $X$, $f_1$, $f_2$, $g$, and $n$, we have

$$\mu_n^T(Xf_1, f_2)(g) + \mu_n^T(f_1, X\,f_2)(g) + \mu_n^T(f_1, f_2)(Xg) = 0. \tag{3.9}$$

Divide by $\|\nu_n\|$ and apply Lemma 2.8 to see:

$$\begin{aligned} &\mu_n^T(ip_n \cdot f_1, f_2)(g) + \mu_n^T(f_1, ip_n \cdot f_2)(g) \\ &= -\frac{\mu_n^T(\mathbf{D}_X f_1, f_2)(g) + \mu_n^T(f_1, \mathbf{D}_X f_2)(g) + \mu_n^T(f_1, f_2)(Xg)}{\|\nu_n\|}, \end{aligned} \tag{3.10}$$

where $p_n(k) = \frac{1}{i}\left\langle \tilde{\nu}_n + \frac{\rho}{\|\nu_n\|}, \mathrm{Ad}(k)X \right\rangle$.

As $n \to \infty$, the right-hand side of (3.10) tends to zero by Lemma 3.5. On the other hand, $p_n f_i$ (considered as continuous functions on $K$) converge uniformly to $p_X f_i$. Another application of Lemma 3.5 shows that the left-hand side of (3.10) converges to $i\mu_\infty^T(p_X f_1, f_2) - i\mu_\infty^T(f_1, p_X \cdot f_2)$. Since $p_X = \overline{p_X}$ this shows that (3.8) holds with $f = p_X$.

Now let $\mathcal{F} \subset C(M\backslash K)$ be the $\mathbb{C}$-subalgebra generated by the $p_X$ and the constant function $1$. Clearly (3.8) holds for all $f \in \mathcal{F}$. This subalgebra is $K$-stable since $p_X(kk_1) = p_{\mathrm{Ad}(k_1)X}(k)$ and hence $\mathcal{F} \cap V_\rho \subset \mathcal{F}$ for all $\rho \in \hat{K}$. Showing $\mathcal{F}$ is dense in $L^2(M\backslash K)$ suffices to conclude that $\mathcal{F} = V_K$.

We will prove the stronger assertion that $\mathcal{F}$ is dense in $C(M\backslash K)$ using the Stone-Weierstrass theorem. Note that $1 \in \mathcal{F}$, and $\mathcal{F}$ is closed under complex conjugation since $p_X = \overline{p_X}$. It therefore suffices to show that $\mathcal{F}$ separates the points of $M\backslash K$. To this end, let $k_1, k_2 \in K$ be such that $p_X(k_1) = p_X(k_2)$ for all $X \in \mathfrak{g}$. Then $\langle \tilde{\nu}_\infty, \mathrm{Ad}(k_1)X\rangle = \langle \tilde{\nu}_\infty, \mathrm{Ad}(k_2)X\rangle$ for all $X \in \mathfrak{g}$, i.e. $\langle \mathrm{Ad}(k_1)^{-1}\tilde{\nu}_\infty - Ad(k_2)^{-1}\tilde{\nu}_\infty, X\rangle = 0$ for all $X \in \mathfrak{g}$. This implies that $\mathrm{Ad}(k_1^{-1})\tilde{\nu}_\infty = \mathrm{Ad}(k_2)^{-1}\tilde{\nu}_\infty$; by the non-degeneracy assumption, $Z_K(\tilde{\nu}_\infty) = Z_K(A) = M$, so $Mk_1 = Mk_2$, i.e. $k_1$ and $k_2$ represent the same point of $M\backslash K$. $\square$

Lemma 3.10 shows easily that $\mu_\infty(\varphi_0, \delta)$ extends to a positive measure. Indeed, choosing $f_j$ as in Lemma 2.5, we see that

$$\mu_\infty(\varphi_0, \delta) = \lim_{j\to\infty} \mu_\infty^T(\varphi_0, |f_j|^2) = \lim_{j\to\infty} \mu_\infty^T(f_j, f_j). \tag{3.11}$$

Here we have invoked Lemma 3.9 for the first equality. It is clear that $\mu_\infty^T(f_j, f_j)$ defines a positive measure on $X$; thus $\mu_\infty(\varphi_0, \delta)$, initially defined as an (algebraic) functional on $C_c^\infty(X)_K$, extends to a positive measure on $X$. To obtain the slightly stronger conclusion implicit in (2) of Theorem 1.6, we will analyze this argument more closely.

**Corollary 3.11.** *Notations as in Lemma 3.10, there exist a constant $C_{f_1,f_2,f}$ and a seminorm $||\cdot||$ on $C_c^\infty(X)$ such that*

$$\left|\mu_n^T(f_1, f\cdot f_2)(g) - \mu_n^T(\overline{f}\cdot f_1, f_2)(g)\right| \le C_{f_1,f_2,f}\|g\| \left[\|\tilde{\nu}_\infty - \tilde{\nu}_n\| + \|\nu_n\|^{-1}\right] \tag{3.12}$$

*Proof.* This follows by keeping track of the error term in the proof of of Lemma 3.10.

Fix a basis $\{X_i\}$ for $\mathfrak{g}$, and define a seminorm on $C_c^\infty(X)$ by $\|g\| = \|g\|_{L^\infty(X)} + \sum_i \|X_i g\|_{L^\infty(X)}$. With this seminorm, (3.12) holds for $f_1, f_2 \in V_K$ and $f = p_X$. This follows from (3.10), utilizing Lemma 3.5 and the fact that $\|p_X - p_n\|_{L^\infty(M\backslash K)} \ll \|\tilde{\nu}_\infty - \tilde{\nu}_n\|$.

Next suppose $f_1, f_2, f, f' \in V_K$ and $\alpha, \alpha' \in \mathbb{C}$. Then, if (3.12) is valid for $(f_1, f_2, f)$ and $(f_1, f_2, f')$, it is also valid for $(f_1, f_2, \alpha f + \alpha' f')$. Further, if (3.12) is valid for $(f_1, f'\cdot f_2, f)$ and for $(\overline{f} f_1, f_2, f')$, then it is also valid for $(f_1, f_2, f\cdot f')$.

Consider now the set of $f \in V_K$ for which (3.12) holds for all $f_1, f_2 \in V_K$. The remarks above show that this is a subalgebra of $V_K$ that contains each $p_X$. The Corollary then follows from the equality $\mathcal{F} = L^2(M\backslash K)_K$ established in the Lemma. $\square$

*Remark* 3.12. In fact, it is possible to obtain a bound of the form $C_{f_1,f_2,f,\tilde{\nu}_n}\|g\|\|\nu_n\|^{-1}$, with the constant uniformly bounded if the $\tilde{\nu}_n$ are uniformly bounded away from the walls. This result can be used to avoid passing to a subsequence in Theorem 1.6 or the following Proposition; this is unnecessary for our applications, however.

**Proposition 3.13.** *(Positivity and equivariance: (2) and (4) of Theorem 1.6).*

*Let $\{\psi_n\}$ be non-degenerate. After replacing $\{\psi_n\}$ by an appropriate subsequence, there exist functions $\tilde{\psi}_n$ on $X$ with the following properties:*

1. *Define the measure $\sigma_n$ via the rule $\sigma_n(g) = \int_X g(x)|\tilde{\psi}_n(x)|^2dx$. Then, for each $g \in C_c^\infty(X)_K$ we have $\lim_{n\to\infty}(\sigma_n(g) - \mu_n(g)) = 0$.*
2. *Let $E \subset \mathrm{End}_G(C^\infty(X))$ be a $\mathbb{C}$-subalgebra of endomorphisms of $C^\infty(X)$, commuting with the $G$-action. Note that each $e \in E$ induces an endomorphism of $C^\infty(Y)$. Assume in addition that $\psi_n$ is an eigenfunction for $E$. Then we may choose $\tilde{\psi}_n$ so that each $\tilde{\psi}_n$ is an eigenfunction for $E$ with the same eigenvalues as $\psi_n$.*

*Proof.* Without loss of generality we may assume that $\{\psi_n\}$ are conveniently arranged.

Let $\{f_j\}_{j=1}^\infty \subset V_K$ be the sequence of functions provided by Lemma 2.5, so that $T(|f_j|^2)$ approximates $\delta$. The main idea is, as in (3.11), to approximate $\mu_n = \mu_n(\varphi_0, \delta)$ using $\mu_n^T(f_j, f_j)$.

For any $g \in C_c^\infty(X)_K$ we have:

$$\begin{aligned}\left|\mu_n(g) - \mu_n^T(f_j, f_j)(g)\right| &\leq \left|\mu_n(\varphi_0, \delta)(g) - \mu_n(\varphi_0, |f_j|^2)(g)\right| \\ &\quad + \left|\mu_n(\varphi_0, |f_j|^2)(g) - \mu_n(f_j, f_j)(g)\right|.\end{aligned} \tag{3.13}$$

Corollary (3.11) provides a seminorm $\|\cdot\|$ on $C_c^\infty(X)$ and a constant $C_j$ such that

$$\left|\mu_n(\varphi_0, |f_j|^2)(g) - \mu_n(f_j, f_j)(g)\right| \leq C_j\|g\| \cdot \left[\|\tilde{\nu}_n - \tilde{\nu}_\infty\| + \|\nu_n\|^{-1}\right].$$

Choose a sequence of integers $\{j_n\}_{n=1}^\infty$ such that $j_n \to \infty$ and:

$$C_{j_n} \cdot \left[\|\tilde{\nu}_n - \tilde{\nu}_\infty\| + \|\nu_n\|^{-1}\right] \xrightarrow[n\to\infty]{} 0$$

We now estimate the other term on the right-hand side of (3.13). Choosing $N = N(g)$ large enough so that $\mu_n(\varphi_0, \delta)(g) = \mu_n(\varphi_0, \delta_N)(g)$, we have

$$\left|\mu_n(\varphi_0, \delta)(g) - \mu_n(\varphi_0, |f_j|^2)(g)\right| \leq \left\||f_j|_N^2 - \delta_N\right\|_{L^2(M\backslash K)}\|g\|_\infty.$$

As $j \to \infty$ (in particular, if $j = j_n$), $|f_j|_N^2 \to \delta_N$ in $V_N$, so this term tends to zero. It follows that

$$\lim_{n\to\infty} \left|\mu_n(g) - \mu_n^T(f_{j_n}, f_{j_n})(g)\right| = 0, \tag{3.14}$$

Setting $\tilde{\psi}_n = R_{\psi_n}(f_{j_n})$, we deduce that

$$\lim_{n\to\infty} \left(\mu_n(g) - \int_X |\tilde{\psi}_n|^2 g(x)dx\right) = 0 \tag{3.15}$$

holds for every $g \in C_c^\infty(X)_K$. In particular, we obtain (1) of the Proposition.

To obtain the equivariance property note that the representation $I_{\nu_n}$ is irreducible as a $(\mathfrak{g}, K)$-module. By [12, Corollary 8.11], there exists $u_n \in U(\mathfrak{g})$ such that $I_{\nu_n}(u_n)\varphi_0 = f_{j_n}$. Thus $\tilde{\psi}_n = u_n\psi_n$. Now every $e \in E$ commutes with the right $G$-action; in particular, $eu_n = u_n e$. It follows that $\tilde{\psi}_n$ transforms under the same character of $E$ as $\psi_n$. □

## 4. Cartan invariance of quantum limits

In this section we show that a nondegenerate quantum limit $\mu_\infty$ is invariant under the action of $A < G$. This invariance follows from differential equations satisfied by the intermediate distributions $\mu_n$. The construction of these differential equations is a purely algebraic problem: construct elements in the $U(\mathfrak{g}_\mathbb{C})$-annihilator of $\varphi_0 \otimes \delta \in V_K \otimes V_K'$, where the $U(\mathfrak{g}_\mathbb{C})$-action is by $I_\nu \otimes I_\nu'$.

Ultimately, these differential equations are derived from the fact that each $z \in \mathfrak{Z} = \mathfrak{Z}(\mathfrak{g}_\mathbb{C})$ acts by a scalar on the representation $(V_K, I_{\nu_n})$. To motivate the method and provide an example, we first work out the simplest case, that of $\mathrm{PSL}_2(\mathbb{R})$, in detail. In this case the resulting operator is due to Zelditch.

### 4.1. Example of $G = \mathrm{PSL}_2(\mathbb{R})$.

Set $G = \mathrm{PSL}_2(\mathbb{R})$, $\Gamma \leq G$ a lattice, and $A$ the subgroup of diagonal matrices. Let $H$ (explicitly given below) be the infinitesimal generator of $A$, thought of first as a differential operator acting on $X = \Gamma\backslash G$ via the differential of the regular representation. If $\{\psi_n\}$ is a conveniently arranged sequence of eigenfunctions on $\Gamma\backslash G/K$, and $\mu_n$ the corresponding distributions (Definition 3.3), we will exhibit a second-order differential operator $J$ such that for all $g \in C_c^\infty(X)_K$,

$$\mu_n((H - \frac{J}{r_n})g) = 0, \tag{4.1}$$

where $r_n \sim |\lambda_n|^{1/2}$. Since the $\mu_n(Jg)$ are bounded (they converge to $\mu_\infty(Jg)$), we will conclude that $\mu_\infty(Hg) = 0$, in other words that $\mu_\infty$ is $A$-invariant. This operator in equation (4.1) is given in [24]. Its discovery was motivated by the proof (via Egorov's theorem) of the invariance of the usual microlocal lift under the geodesic flow. We show here how it arises naturally in the representation-theoretic approach.

By Lemma 3.7, it will suffice to find an operator annihilating the element $\varphi_0 \otimes \delta \in V_K \otimes V_K'$, where $U(\mathfrak{g}_\mathbb{C})$ acts via $I_\nu \otimes I_\nu'$.

Let $H = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$, $X_+ = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$, $X_- = \begin{pmatrix} 0 & \\ 1 & 0 \end{pmatrix}$ be the standard generators of $\mathfrak{sl}_2$, with the commutation relations $[H, X_\pm] = \pm 2X_\pm$, $[X_+, X_-] = H$. The roots w.r.t. the maximal split torus $\mathfrak{a} = \mathbb{R} \cdot H$ are given by $\pm\alpha(H) = \pm 2$. We also set $W = X_+ - X_-$, so that $\mathbb{R} \cdot W = \mathfrak{k}$. Letting $+\alpha$ be the positive root, $\mathfrak{n} = \mathbb{R} \cdot X_+$, we have $\rho(H) = \frac{1}{2}\alpha(H) = 1$. Set $\exp \mathfrak{a} = A$ as in the introduction.

The Casimir element $C \in \mathfrak{Z}(\mathfrak{sl}_2\mathbb{C})$ is given by $4C = H^2 + 2X_+X_- + 2X_-X_+$. For the parameter $\nu \in i\mathfrak{a}^*$ given by $\nu(H) = 2ir$ ($r \in \mathbb{R}$), $C$ acts on $\pi_\nu$ with the eigenvalue $\lambda = -\frac{1}{4} - r^2$. The Weyl element acts by mapping $\nu \mapsto -\nu$. On $S = G/K$ with the metric normalized to have constant curvature $-1$, $C$ reduces to the hyperbolic Laplacian. In particular, every eigenfunction $\psi \in L^2(\Gamma\backslash G/K)$ with eigenvalue $\lambda < -\frac{1}{4}$ generates a unitary principal series subrepresentation. Definition 3.3 associates to $\psi$ a distribution $\mu_\psi(\varphi_0, \delta)$ on $\Gamma\backslash G$.

As in Definition 2.6, we have an action $I_\nu$ of $G$ on $V$ and of $\mathfrak{g}$ on $V_K$. Note that for $g \in NA$, $f \in V_K$, $(I_\nu(g)f)(1) = f(g) = e^{\langle \nu+\rho, H_0(g)\rangle} f(1)$. Since $\delta(f) = f(1)$ and the pairing between $V_K$ and $V_K'$ is $G$-invariant, it follows that for $X \in \mathfrak{a} \oplus \mathfrak{n}$, $I_\nu'(X)\delta = -\langle \nu + \rho, H_0(X)\rangle \delta$.

Suppressing $I_\nu$ from now on this means that $X \cdot (f \otimes \delta) = (Xf) \otimes \delta - \langle \nu + \rho, H_0(X)\rangle f \otimes \delta$. Extend $\nu + \rho$ trivially on $\mathfrak{n}$ to obtain a functional on $\mathfrak{a} \oplus \mathfrak{n}$. Then

$$(X + (\nu + \rho)(X)) \cdot (f \otimes \delta) = (Xf) \otimes \delta. \tag{4.2}$$

Now since $\mathfrak{a}$ normalizes $\mathfrak{n}$ and $\nu + \rho$ is trivial on $\mathfrak{n}$, the map $X \mapsto X + (\nu + \rho)(X)$ is a Lie algebra homomorphism $\mathfrak{a} \oplus \mathfrak{n} \to \mathfrak{a} \oplus \mathfrak{n}$, and hence extends to an algebra homomorphism $\tau_{\nu+\rho} \colon U(\mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C}) \to U(\mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C})$. (4.2) shows that, for $u \in U(\mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C})$,

$$\tau_{\nu+\rho}(u) \cdot (f \otimes \delta) = (uf) \otimes \delta \tag{4.3}$$

In view of (4.3) any operator $u \in U(\mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C})$ annihilating $\varphi_0$ gives rise to an operator annihilating $\varphi_0 \otimes \delta$.

The natural starting point is the eigenvalue equation $(4C+1+4r^2)\varphi_0 = 0$. Of course, $C$ is not an element of $U(\mathfrak{n}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C})$. Fortunately, it "nearly" is: there exists an $C' \in U(\mathfrak{n}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C})$ such that $C - C'$ annihilates $\varphi_0$.

In detail, we use the commutation relations and the fact that $X_- = X_+ - W$ to write $4C = H^2 - 2H + 4X_+^2 - 4X_+W$. Since $\phi_0$ is spherical, it follows that $W\phi_0 = 0$. Thus

$$\left(H^2 - 2H + 4X_+^2 + 1 + 4r^2\right)\varphi_0 = 0 \tag{4.4}$$

Since $(\nu+\rho)(H) = 2ir + 1$, we conclude from (4.3) that:

$$\left((H + 2ir + 1)^2 - 2(H + 2ir + 1) + 4X_+^2 + 1 + 4r^2\right)\cdot\varphi_0 \otimes \delta = 0.$$

Collecting terms in powers of $r$ we see that this may be written as:

$$\left((2H)(2ir) + (H^2 + 4X_+^2)\right)\varphi_0 \otimes \delta = 0$$

Setting $J = \frac{H^2+4X_+^2}{4i}$ and dividing by $4ir$ we see that the operator $H + \frac{J}{r}$ annihilates $\varphi_0 \otimes \delta$, and so also the distribution $\mu_n$. One then deduces the $A$-invariance of $\mu_\infty$ as discussed in the start of this section.

Notice that the terms involving $r^2$ in (4.4) canceled. This is a general feature which will be of importance.

4.2. **The general proof.** We now generalize these steps in order. Notations being as in Section 2 and Definition 3.3, we first compute the action of $U(\mathfrak{m}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C})$ on $\delta$ (Lemma 4.1) and then on $\varphi_0 \otimes \delta$ (Corollary 4.2). Secondly we find an appropriate form for the elements of $\mathfrak{Z}(\mathfrak{g}_\mathbb{C})$ (Corollary 4.4), which gives us the exact differential equation (4.6). We then show that the elements we constructed annihilating $\mu_\psi$ are (up to scaling) of an appropriate form $H + \frac{J}{\|\nu\|^*}$ (Lemma 4.5), and "take the limit as $\nu \to \infty$" (Corollary 4.6) to see that $\mu_\infty$ is invariant under a sub-torus of $A$.

A final step (not so apparent in the $\mathrm{PSL}_2(\mathbb{R})$ case) is to verify that we have constructed *enough* differential operators to obtain invariance under the full split torus (Lemma 4.7). In fact, even in the rank-1 case one needs to verify that the "$H$" part is non-zero.

Given $\lambda \in \mathfrak{a}_\mathbb{C}^*$, we extend it to a linear map $\mathfrak{m}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C} \to \mathbb{C}$. Since $\mathfrak{m}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C}$ is an ideal of this Lie algebra, $\lambda$ is a Lie algebra homomorphism; thus it extends to an algebra homomorphism $\lambda\colon U(\mathfrak{m}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C}) \to \mathbb{C}$. We denote by $\tau_\lambda$ the translation automorphism of $U(\mathfrak{m}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C})$ given by $X \mapsto X + \lambda(X)$ on $\mathfrak{m}_\mathbb{C} \oplus \mathfrak{a}_\mathbb{C} \oplus \mathfrak{n}_\mathbb{C}$. Similarly, given $\chi \in \mathfrak{h}_\mathbb{C}^*$, we define $\tau_\chi : U(\mathfrak{h}_\mathbb{C}) \to U(\mathfrak{h}_\mathbb{C})$. We shall write $U(\mathfrak{g}_\mathbb{C})^{\le d}$ for the elements of $U(\mathfrak{g}_\mathbb{C})$ of degree $\le d$, and similarly for other enveloping algebras and $\mathfrak{Z} = \mathfrak{Z}(\mathfrak{g}_\mathbb{C})$ (e.g. $\mathfrak{Z}^{\le d} = \mathfrak{Z} \cap U(\mathfrak{g}_\mathbb{C})^{\le d}$).

Let $\nu \in \mathfrak{a}_\mathbb{C}^*$. Let $\chi_\nu : \mathfrak{Z} \to \mathbb{C}$ be the infinitesimal character corresponding to $I_\nu$ (that is, the scalar by which $\mathfrak{Z}$ acts in $(I_\nu, V_K)$.) Recall that $\rho_\mathfrak{h}$ denotes the half-sum of positive roots for $(\mathfrak{h}_\mathbb{C} : \mathfrak{g}_\mathbb{C})$, $\rho$ the half-sum for $(\mathfrak{a} : \mathfrak{g})$.

**Lemma 4.1.** *For $X \in \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}$, $I_\nu(X)\delta = -\langle \nu + \rho, X\rangle\,\delta$.*

*Proof.* This follows from the definitions. $\square$

**Corollary 4.2.** *For any* $u \in U(\mathfrak{m}_{\mathbb{C}} \oplus \mathfrak{a}_{\mathbb{C}} \oplus \mathfrak{n}_{\mathbb{C}})$ *and* $f \in V_K$,

$$I_\nu \otimes I'_\nu(\tau_{\nu+\rho}(u)) \cdot (f \otimes \delta) = (I_\nu(u)f) \otimes \delta.$$

*Proof.* This follows from the previous Lemma. □

**Definition.** Let $\mathrm{pr} : U(\mathfrak{g}_{\mathbb{C}}) \to U(\mathfrak{h}_{\mathbb{C}})$ be the projection corresponding to the decomposition $U(\mathfrak{g}_{\mathbb{C}}) = U(\mathfrak{h}_{\mathbb{C}}) \oplus [(\mathfrak{n}_{\mathbb{C}} \oplus \mathfrak{n}_M)U(\mathfrak{g}_{\mathbb{C}}) + U(\mathfrak{g}_{\mathbb{C}})(\bar{\mathfrak{n}}_{\mathbb{C}} \oplus \bar{\mathfrak{n}}_M])$ (arising from the decomposition $\mathfrak{g}_{\mathbb{C}} = \mathfrak{n}_{\mathbb{C}} \oplus \mathfrak{n}_M \oplus \mathfrak{h}_{\mathbb{C}} \oplus \bar{\mathfrak{n}}_{\mathbb{C}} \oplus \bar{\mathfrak{n}}_M$ by the Poincaré-Birkhoff-Witt Theorem).

**Lemma 4.3.** *For* $z \in \mathfrak{Z}^{\leq d}$, *we have*

$$z - \mathrm{pr}(z) \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})^{\leq d-2}U(\mathfrak{k}_{\mathbb{C}}).$$

*Proof.* It suffices to show that $z - \mathrm{pr}(z) \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}U(\mathfrak{k}_{\mathbb{C}})$, since $\mathfrak{g}_{\mathbb{C}} = \mathfrak{n}_{\mathbb{C}} \oplus \mathfrak{a}_{\mathbb{C}} \oplus \mathfrak{k}_{\mathbb{C}}$.

Let $\mathcal{B}(\mathfrak{n}_{\mathbb{C}})$, $\mathcal{B}(\bar{\mathfrak{n}}_{\mathbb{C}})$, $\mathcal{B}(\mathfrak{n}_M)$ and $\mathcal{B}(\bar{\mathfrak{n}}_M)$ be bases for $\mathfrak{n}_{\mathbb{C}}$, $\bar{\mathfrak{n}}_{\mathbb{C}}$, $\mathfrak{n}_M$ and $\bar{\mathfrak{n}}_M$, respectively, consisting of $\mathfrak{h}_{\mathbb{C}}$-eigenvectors. Let $\mathcal{B}(\mathfrak{a}_{\mathbb{C}})$ and $\mathcal{B}(\mathfrak{b}_{\mathbb{C}})$ be bases for $\mathfrak{a}_{\mathbb{C}}$ and $\mathfrak{b}_{\mathbb{C}}$, respectively.

By Poincaré-Birkhoff-Witt, one may uniquely express $z$ as a linear combination of terms of the form:

$$\mathcal{D} = X_1 \dots X_n Y_1 \dots Y_m A_1 \dots A_t B_1 \dots B_r \overline{X}_1 \dots \overline{X}_k \overline{Y}_1 \dots \overline{Y}_l$$

where $X_* \in \mathcal{B}(\mathfrak{n}_{\mathbb{C}})$, $Y_* \in \mathcal{B}(\mathfrak{n}_M)$, $A_* \in \mathcal{B}(\mathfrak{a}_{\mathbb{C}})$, $B_* \in \mathcal{B}(\mathfrak{b}_{\mathbb{C}})$,$\overline{X}_* \in \mathcal{B}(\bar{\mathfrak{n}}_{\mathbb{C}})$ and $\overline{Y}_* \in \mathcal{B}(\bar{\mathfrak{n}}_M)$. Then $z - \mathrm{pr}(z)$ consists of the sum of all terms $\mathcal{D}$ for which $n + m + k + l \neq 0$. We show that each such term satisfies $\mathcal{D} \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}U(\mathfrak{k}_{\mathbb{C}})$.

In view of the fact that $z - \mathrm{pr}(z)$ commutes with $\mathfrak{a}_{\mathbb{C}}$, one has $n = 0$ iff $k = 0$. Further, if $n = k = 0$, then the fact that $z - \mathrm{pr}(z)$ commutes with $\mathfrak{b}_{\mathbb{C}}$ implies $m = 0$ iff $l = 0$. Also one has $n + m + t + r + k + l \leq d$.

We now proceed in a case-by-case basis, using either the inclusion $\mathfrak{n}_M \oplus \mathfrak{b}_{\mathbb{C}} \oplus \bar{\mathfrak{n}}_M = \mathfrak{m}_{\mathbb{C}} \subset \mathfrak{k}_{\mathbb{C}}$, or the observation that for $\overline{X} \in \bar{\mathfrak{n}}_{\mathbb{C}}$ we have $\theta_{\mathbb{C}}\overline{X} \in \mathfrak{n}_{\mathbb{C}}$, while $\overline{X} + \theta_{\mathbb{C}}\overline{X} \in \mathfrak{k}_{\mathbb{C}}$ (it is $\theta_{\mathbb{C}}$-stable!).

(1) $k = l = 0$ is impossible, for this would force $n = m = 0$.
(2) $k \geq 1$ and $l \geq 1$. Then $n \geq 1$ so that $X_1 \dots X_n \in U(\mathfrak{n}_{\mathbb{C}})$, $\overline{Y}_1 \dots \overline{Y}_l \in U(\mathfrak{k}_{\mathbb{C}})$, and $m + t + r + k \leq d - 2$.
(3) $k = 0$ and $l \geq 1$. Then $n = 0$ and $m \geq 1$, so $t \leq d - 2$. Since $[\mathfrak{a}, \mathfrak{m}] = 0$ we may commute the $A$-terms past the $Y$-terms, so that $\mathcal{D}$ is the product of the $A$-terms (at most $d - 2$ of them) and $Y_1 \dots Y_m B_1 \dots B_r \overline{Y}_1 \dots \overline{Y}_l \in U(\mathfrak{k}_{\mathbb{C}})$.
(4) $k \geq 1$ and $l = 0$. Then $n \geq 1$. Set $s = Y_1 \dots Y_m A_1 \dots A_t B_1 \dots B_r \overline{X}_1 \dots \overline{X}_{k-1}$ so that $\mathcal{D} = X_1 \dots X_n \cdot s \cdot \overline{X}_k$. Since $m+t+r+(k-1) \leq d-1-n \leq d-2$, we have $s \in U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}$. Then (recall $\theta_{\mathbb{C}}$ is the complex-linear

extension of the Cartan involution $\theta$ to $\mathfrak{g}_{\mathbb{C}}$),

$$\begin{aligned}(4.5)\qquad \mathcal{D} = X_1 \dots X_n s \overline{X}_k &= X_1 \dots X_n \cdot s \cdot (\overline{X}_k - \theta_{\mathbb{C}}(\overline{X}_k)) \\ &+ X_1 \dots X_n \theta_{\mathbb{C}}(\overline{X}_k) s \\ &+ X_1 \dots X_n (s\theta_{\mathbb{C}}(\overline{X}_k) - \theta_{\mathbb{C}}(\overline{X}_k) s).\end{aligned}$$

From the observation above, the first two terms on the right clearly belong to $U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}U(\mathfrak{k}_{\mathbb{C}})$. Moreover, $\left[s, \theta_{\mathbb{C}}(\overline{X}_k)\right] \in U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}$ (for any $p \in U(\mathfrak{g}_{\mathbb{C}})^{\leq d_p}, q \in U(\mathfrak{g}_{\mathbb{C}})^{\leq d_q}$ the general fact $[p, q] \in U(\mathfrak{g}_{\mathbb{C}})^{d_p+d_q-1}$ follows by induction on the degrees from the formula $[ab, c] = a[b, c] + [a, c]b$). Thus the third term of (4.5) belongs to $U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{g}_{\mathbb{C}})^{\leq d-2}U(\mathfrak{k}_{\mathbb{C}})$ also.

□

**Corollary 4.4.** *Let $z \in \mathfrak{Z}^{\leq d}$. Then there exists $b = b(z) \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})^{\leq d-2}$ such that $z - \mathrm{pr}(z) + b(z) \in U(\mathfrak{g}_{\mathbb{C}}) \cdot \mathfrak{k}_{\mathbb{C}}$.*

Since $I_\nu(\mathfrak{k}_{\mathbb{C}})$ annihilates $\varphi_0$ and $z \cdot \varphi_0 = \chi_\nu(z)\varphi_0$, we have $I_\nu(\chi_\nu(z) - \mathrm{pr}(z) + b(z)) \cdot \varphi_0 = 0$. In view of Corollary 4.2 we obtain:

$$(4.6)\qquad I_\nu \otimes I'_\nu(\tau_{\nu+\rho}\,\mathrm{pr}(z) - \tau_{\nu+\rho} b(z) - \chi_\nu(z))(\varphi_0 \otimes \delta) = 0$$

In what follows, we shall freely identify the algebra $U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$ with the Weyl-invariant polynomial functions on $\mathfrak{h}^*_{\mathbb{C}}$.

Given $\mathcal{P} \in U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$, we denote by $\mathcal{P}' \colon \mathfrak{h}^*_{\mathbb{C}} \to \mathfrak{h}_{\mathbb{C}}$ its differential. In other words, we identify $\mathcal{P}$ with a polynomial function on $\mathfrak{h}^*_{\mathbb{C}}$, and $\mathcal{P}'$ denotes the derivative of this function; it takes values in the cotangent space of $\mathfrak{h}^*_{\mathbb{C}}$, which is canonically identified at every point with $\mathfrak{h}_{\mathbb{C}}$.

We shall use the notation $U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq r}$ to denote polynomials of degree $\leq r$ on $\mathfrak{a}^*_{\mathbb{C}}$, valued in the vector space $U(\mathfrak{g}_{\mathbb{C}})$. Note that given $J \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq r}$ and $\nu \in \mathfrak{a}^*_{\mathbb{C}}$ we can speak of the "value of $J$ at $\nu$." We denote it by $J(\nu)$ and it belongs to $U(\mathfrak{g}_{\mathbb{C}})$.

**Lemma 4.5.** *Let $\mathcal{P} \in U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$ have degree $\leq d$. Set $H = \frac{\mathcal{P}'(\nu)}{||\nu||^{d-1}} \in \mathfrak{h}_{\mathbb{C}}$. Then there exists $J \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq d-2}$ such that*

$$I_\nu \otimes I'_\nu \left( H + \frac{J(\nu)}{\|\nu\|^{d-1}} \right) \cdot \varphi_0 \otimes \delta = 0.$$

*(As defined in Section 2, $\|\nu\|$ denotes the norm of $\nu \in \mathfrak{a}^*_{\mathbb{C}}$ w.r.t. the Killing form.)*

*Proof.* The map $\gamma_{HC} \colon \mathfrak{Z} \to U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$ given by $\gamma_{HC}(z) = \tau_{\rho_{\mathfrak{h}}}\,\mathrm{pr}(z)$ is an isomorphism of algebras, the Harish-Chandra homomorphism. With the above identification, the infinitesimal character of $(V_K, I_\nu)$ corresponds to "evaluation at $\nu + \rho - \rho_{\mathfrak{h}}$," i.e. for $\mathcal{P} \in U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$:

$$\chi_\nu(\gamma_{HC}^{-1}(\mathcal{P})) = \mathcal{P}(\nu + \rho_{\mathfrak{a}} - \rho_{\mathfrak{h}}) \tag{4.7}$$

(See [12, Prop 8.22]; w.r.t. the maximal torus $\mathfrak{b}_{\mathbb{C}} \subset \mathfrak{m}_{\mathbb{C}}$, the infinitesimal character of the trivial representation of $\mathfrak{m}_{\mathbb{C}}$ is (the Weyl-group orbit of) $\rho - \rho_{\mathfrak{h}}$).

Given $\mathcal{P} \in U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$ of degree $d$, we set $z = \gamma_{HC}^{-1}(\mathcal{P})$ in (4.6), writing $b(\mathcal{P})$ for the element $b(z)$. Note that $z \in Z(\mathfrak{g}_{\mathbb{C}})^{\leq d}$, as the Harish-Chandra homomorphism "preserves degree" (see [3, 7.4.5(c)]), and hence $b(\mathcal{P}) \in U(\mathfrak{n}_{\mathbb{C}})U(\mathfrak{a}_{\mathbb{C}})^{\leq d-2}$.

Combining (4.6) and (4.7): $\varphi_0 \otimes \delta$ is then annihilated by the operator

$$\left(\tau_{\nu+\rho-\rho_{\mathfrak{h}}}\mathcal{P} - \mathcal{P}(\nu + \rho_{\mathfrak{a}} - \rho_{\mathfrak{h}}) - \tau_{\nu+\rho} b(\mathcal{P})\right) \varphi_0 \otimes \delta = 0 \tag{4.8}$$

Let $\underline{x} = (x_1, \ldots, x_n), \underline{y} = (y_1, \ldots, y_n)$. If a polynomial $p \in \mathbb{C}[\underline{x}]$ has degree $d$, $p(\underline{x} + \underline{y}) - p(\underline{y}) = p'(\underline{y})(\underline{x}) + q(\underline{x}, \underline{y})$ where $q \in (\mathbb{C}[\underline{x}])\,[\underline{y}]$ has degree at most $d-2$ in $\underline{y}$, and the derivative $p'(\underline{y})$ is understood to act as a linear functional on $\underline{x}$.

Applying this to $p = \mathcal{P}, \underline{y} = \nu + \rho - \rho_{\mathfrak{h}}$ we see that there exists $J_1 \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq d-2}$ with $\deg(J) \leq d-2$ and

$$\tau_{\nu+\rho-\rho_{\mathfrak{h}}}\mathcal{P} - \mathcal{P}(\nu + \rho - \rho_{\mathfrak{h}}) = \mathcal{P}'(\nu + \rho - \rho_{\mathfrak{h}}) + J_1(\nu) \tag{4.9}$$

Now $b(\mathcal{P}) \in U(\mathfrak{n}_{\mathbb{C}}) \cdot U(\mathfrak{a}_{\mathbb{C}})^{\leq d-2}$, so the map $\nu \mapsto \tau_{\nu+\rho} b(\mathcal{P})$ can be regarded as an element $J_2 \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq d-2}$. Similarly $\nu \mapsto \mathcal{P}'(\nu + \rho - \rho_{\mathfrak{h}}) - \mathcal{P}'(\nu)$ defines an element $J_3 \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq d-2}$.

Combining these remarks with (4.8) and (4.9), we see that

$$(\mathcal{P}'(\nu) + J_1(\nu) + J_2(\nu) + J_3(\nu))\varphi_0 \otimes \delta = 0$$

Set $J \equiv J_1 + J_2 + J_3$ and divide by $\|\nu\|^{d-1}$ to conclude. □

**Corollary 4.6.** *Let $\mathcal{P} \in U(\mathfrak{h}_{\mathbb{C}})^{W(\mathfrak{h}_{\mathbb{C}}:\mathfrak{g}_{\mathbb{C}})}$. Notations being as in Definition 3.8 and Lemma 3.9, suppose $\{\psi_n\}$ is conveniently arranged. Then $\mu_\infty(\varphi_0, \delta)$ is $\mathcal{P}'(\tilde{\nu}_\infty)$-invariant.*

*Proof.* It suffices to verify this for $\mathcal{P}$ homogeneous, say of degree $d$. Combining Lemma 4.5 and Lemma 3.7, and using the homogeneity of $\mathcal{P}$, we see that there exists $J \in U(\mathfrak{g}_{\mathbb{C}})[\mathfrak{a}_{\mathbb{C}}]^{\leq d-2}$ so that

$$\left(\mathcal{P}'(\tilde{\nu}_n) + \frac{J(\nu_n)}{||\nu_n||^{d-1}}\right) \cdot \mu_n(\varphi_0 \otimes \delta) = 0$$

Here $(\mathcal{P}' + \ldots)$ acts on $\mu_n(\varphi_0 \otimes \delta)$ according to the natural action of $U(\mathfrak{g}_{\mathbb{C}})$ on $C_c^\infty(X)'_K$. Now fix $g \in C_c^\infty(X)_K$. Let $u \to u^t$ be the unique $\mathbb{C}$-linear anti-involution of $U(\mathfrak{g}_{\mathbb{C}})$ such that $X^t = -X$ for $X \in \mathfrak{g}_{\mathbb{C}} \subset U(\mathfrak{g}_{\mathbb{C}})$.

Then we have for each $d$

$$\mu_n\left(\varphi_0 \otimes \delta\right)\left(\left(\mathcal{P}'(\tilde{\nu}_n) - \frac{J^t(\nu_n)}{\|\nu_n\|^{d-1}}\right) g\right) = 0. \tag{4.10}$$

Note that, as $n$ varies, the quantity $\left(\mathcal{P}'(\tilde{\nu}_n) - \frac{J^t(\nu_n)}{\|\nu_n\|^{d-1}}\right) g$ remains in a fixed finite dimensional subspace of $C^\infty(X)_K$. Further, it converges in that subspace to $\mathcal{P}'(\tilde{\nu}_\infty)g$.

With these remarks in mind, we can pass to the limit $n \to \infty$ in (4.10) to obtain $\mu_\infty(\varphi_0 \otimes \delta)(\mathcal{P}'(\nu_\infty)g) = 0$, i.e. $\mathcal{P}'(\nu_\infty)$ annihilates $\mu_\infty$ as required. $\square$

It remains to show that the subspace

$$S = \left\{\mathcal{P}'(\tilde{\nu}_\infty) \mid \mathcal{P} \in U(\mathfrak{h}_\mathbb{C})^{W(\mathfrak{h}_\mathbb{C}\,:\,\mathfrak{g}_\mathbb{C})}\right\} \subset \mathfrak{h}_\mathbb{C} \tag{4.11}$$

contains $\mathfrak{a}_\mathbb{C}$. By the Corollary this will show that $\mathfrak{a}$ annihilates any limit measure, or that this measure is $A$-invariant.

**Lemma 4.7.** *Let $W_0 \subset W(\mathfrak{h}_\mathbb{C}\,:\,\mathfrak{g}_\mathbb{C})$ be the stabilizer of $\tilde{\nu}_\infty \in \mathfrak{a}_\mathbb{C}^*$, and define $S$ as in (4.11). Then $S = \mathfrak{h}_\mathbb{C}^{W_0}$. In particular, if $\tilde{\nu}_\infty$ is regular, then $S$ contains $\mathfrak{a}_\mathbb{C}$.*

*Proof.* This can be seen either from the fact that $S$ is the image of the map on cotangent spaces induced by the quotient map $\mathfrak{h}_\mathbb{C}^* \to \mathfrak{h}_\mathbb{C}^*/W_0$, or more explicitly: first construct many elements in $U(\mathfrak{h}_\mathbb{C})^{W(\mathfrak{h}_\mathbb{C}\,:\,\mathfrak{g}_\mathbb{C})}$ by averaging over $W(\mathfrak{h}_\mathbb{C}\,:\,\mathfrak{g}_\mathbb{C})$, and then directly compute derivatives to obtain the claimed equality.

$W_0$ is generated by the reflections in $W(\mathfrak{h}_\mathbb{C}\,:\,\mathfrak{g}_\mathbb{C})$ fixing $\tilde{\nu}_\infty$. In the case where $\tilde{\nu}_\infty$ is regular as an element in $i\mathfrak{a}_\mathbb{R}^*$, the corresponding roots must be trivial on all of $\mathfrak{a}_\mathbb{C}^*$. In particular, any element of $W_0$ fixes all of $\mathfrak{a}_\mathbb{C}$. $\square$

**Corollary 4.8.** *Let notations be as in Proposition 3.13. Then any weak-* limit $\sigma_\infty$ of the measures $\sigma_n$ is $A$-invariant.*

*Proof.* After passing to an appropriate subsequence, we may assume that $\{\psi_n\}$ is conveniently arranged. Proposition 3.13, (1), shows that $\sigma_\infty(g) = \mu_\infty(\varphi_0, \delta)(g)$ whenever $g \in C_\mathrm{c}^\infty(X)_K$. Corollary 4.6 and Lemma 4.7, together with the fact that $C_\mathrm{c}^\infty(X)_K$ is dense in $C_0(X)$, show that $\sigma_\infty$ is $A$-invariant. $\square$

## 5. Complements

In this section we gather together several points complementing the main text.

5.1. **Extensions to general $G$.** In practice we wish to apply our result to groups which are slightly more general that the ones considered above. Here we briefly discuss extensions of the present work to reductive groups.

From now on let $G$ be a linear connected reductive Lie group, $\Gamma < G$ a lattice. (For "linear connected reductive", we follow the definition of [12, Chapter 1].) We set $X = \Gamma\backslash G$ as before, and define in addition $X_Z = Z\Gamma\backslash G$, with $Z = Z(G)$. $X_Z$ has finite volume w.r.t. the $G$-invariant measure. In a similar fashion we shall consider $Y = X/K$ and $Y_Z = X_Z/K$.

Since $G$ is linear connected reductive, we have a decomposition $\mathfrak{g} = \mathfrak{z} \oplus_j \mathfrak{g}^{(j)}$ where $\mathfrak{z} = Z_{\mathfrak{g}}$, and each $\mathfrak{g}^{(j)}$ is a simple Lie algebra (an orthogonal decomposition w.r.t. the Killing form), leading to a decomposition $G = Z_G \times \prod_j G^{(j)}$ (almost direct product), where the $G^{(j)}$ are connected semisimple or compact normal subgroups.

Choosing the Cartan involution, the subalgebra $\mathfrak{a}$, etc. compatible with this decomposition, let $K = K_Z \times \prod_j K^{(j)}$ be the $\Theta$-fixed maximal compact subgroup. If $G^{(j)}$ is compact then $K^{(j)} = G^{(j)}$, of course. Note that the subgroup $M = Z_K(\mathfrak{a})$ now includes the compact part of the center, as well as all compact factors.

For a unitary character $\omega \in \hat{Z}$, let $L^2(X, \omega)$ denote the space of all measurable $f : X \to \mathbb{C}$ such that $f(zg) = \omega(z)f(g)$ for all $z \in Z$, and such that $\|f\|^2 = \int_{X_Z} |f(x)|^2 dx < \infty$. If $\omega$ is *unramified* (i.e. trivial on $Z(G) \cap K$), then set $L^2(Y, \omega) = L^2(X, \omega)^K$. If $\omega$ is unramified and $\psi \in L^2(Y, \omega)$, then $|\psi(y)|^2$ is $Z$-invariant, and we can define a finite measure $\bar{\mu}_\psi$ on $Y_Z$ as before.

An eigenfunction $\psi \in L^2(Y, \omega)$ still generates an irreducible subrepresentation of $G$ in $L^2(X, \omega)$. From this we obtain, as in Section 3, a norm-reducing intertwining operator $R_\psi \colon (V_K, I_\nu) \to L^2(X, \omega)$, and (as in Definition 3.1) a map $\mu_\psi \colon V_K \otimes V_K' \to (C_c^\infty(X_Z)_K)'$ as before (note that for $f_1, f_2 \in V_K$, $R_\psi(f_1)\overline{R_\psi(f_2)}$ is $Z$-invariant since $\omega$ is unitary, and as before its $L^1$ norm is at most the product of the $L^2$ norms of $f_1, f_2 \in V$).

Let $\{\omega_n\}_{n=1}^\infty$ be a sequence of unramified characters of $Z$. We now consider a sequence of eigenfunctions $\{\psi_n\}_{n=1}^\infty$ such that $\psi_n \in L^2(Y, \omega_n)$, with intertwining operators $R_n$ and parameters $\nu_n \in \mathfrak{a}_{\mathbb{C}}^*$, and assume that the $\nu_n$ escape to infinity.

**Definition 5.1.** *Call the sequence* non-degenerate *if for every non-compact $j$, the sequence $\{\nu_n^{(j)}\} \subset \left(\mathfrak{a}_{\mathbb{C}}^{(j)}\right)^*$ is non-degenerate in the sense of Definition 3.8.*

*Remark* 5.2. As before, for a non-degenerate sequence we have $\mathrm{Re}(\nu_n) = 0$ and $\mathrm{Im}(\nu_n)$ regular for large enough $n$. However, the rates at which the different components of $\nu_n$ tend to infinity need not be the same.

Indeed, defining a $\tilde{\nu}_n^{(j)}$ for each $j$ by normalizing $\nu_n^{(j)}$, and passing to a subsequence where they all converge, the non-degeneracy assumption amounts to assuming that the limits $\tilde{\nu}^{(j)}$ are regular (i.e. do not lie on any wall). Of course, the rate of convergence at different $j$ may be different.

Lemma 3.10 and its subsequent Corollary continue to hold (replace $C_c^\infty(X)$ with $C_c^\infty(X_Z)$). The only modification to the proof is that one should only consider functions $p_X$ given by $X \in \mathfrak{g}^{(j)}$, rescaling by $\left\|\nu_n^{(j)}\right\|$. The Stone-Weierstrass argument will show that the algebra generated by these "limited" $p_X$ is dense. Defining the lift as before (using the $\delta$ distribution at $1 \in M\backslash K$), we obtain the positivity of the limits.

In the same vein it is clear that by using $U(\mathfrak{g}_{\mathbb{C}}^{(j)})$ and its center (which is contained in the center of $U(\mathfrak{g}_{\mathbb{C}})$), the analysis of Section 4 shows that a non-degenerate limit is $\mathfrak{a}^{(j)}$-invariant for all non-compact $j$, and hence $A$-invariant. As before, every $\mu_{R_n}$ is $M$-invariant, hence so is $\mu_\infty$.

5.2. **Degenerate limits.** It is an interesting and natural problem to extend the results of the present paper to degenerate limits, i.e. sequence of eigenfunctions $\psi_n$ such that $\frac{\nu_n}{\|\nu_n\|}$ converges to one of the walls of a Weyl chamber.

The non-degeneracy assumption was used in several places in the above arguments. The first was in the assertion that the intertwining maps from the models $(\pi_\nu, V_K)$ to $L^2(X)$ were isometries for the $L^2$ norm on $V_K$, so that the total variation of the measures $\mu_\psi^T(f_1, f_2)$ was bounded independently of the parameter $\nu$ of $\psi$. Secondly, we used it in the proof of positivity of the limit measures by integration by parts. Finally, it was used to conclude that the limit measures is indeed invariant under the full Cartan subgroup $A$.

The first use can be removed in a straightforward manner resulting in a lift of the limit measure which is a positive measure on $X$. However, the question of invariance is more subtle, and one might expect the methods presented here to only show invariance under an appropriate subtorus of $A$. We hope to revisit this issue in the future.

5.3. **Geometry of the Cartan flow and flats.** A symmetric space comes with a rich structure of flat subspaces; these are an important part of the large-scale geometry of the space. Our aim here is to discuss the connection of the Cartan flow (i.e. the action of $A$ on $X/M$) with the structure of flats. Crudely speaking, the Cartan flow is analogous to the geodesic flow, but with "geodesic" replaced by "flat." This highlights the fact that the present result is a generalization of the rank 1 situation, where flats *are* geodesics. (The present result, however, is new even in the case of hyperbolic 3-space, on account of its equivariance.)

Let $G$ and other notations be as fixed in Section 2.1, and let $r$ be the real rank of $G$, $W = W(\mathfrak{a} : \mathfrak{g})$ the Weyl group.

An *$r$-flat* in $S$ is, by definition, a subspace isometric to $\mathbb{R}^r$ with the flat metric. Given any $r$-flat $F \subset S$ and a point $P \in F$, there is a canonical $W$-conjugacy class $\mathcal{C}_{P,F}$ of isometries from $\mathfrak{a}$ to $F$, all mapping $0$ to $P$. Indeed, we may assume that $P = x_K$, in which case we may identify $F$ (via the inverse exponential mapping) with a subset of $\mathfrak{p}$, which may be shown (see [16]) to be a maximal abelian subspace. In particular, this subset is conjugate under $K$ to $\mathfrak{a}$, and this conjugacy is unique up to the action of the Weyl group, whence the assertion.

An *orientation* $\varphi$ of the pair $(P, F)$ will be an element $\varphi \in \mathcal{C}_{P,F}$; there are therefore precisely $|W|$ orientations for any pair $(P, F)$. A *chamber* will be a triple $(P, F, \varphi)$ of a point $P$, a flat $F$ containing $P$, and an orientation for $(P, F, \varphi)$. In the case $r = 1$, a chamber is equivalent to a geodesic ray: given a chamber $(P, F, \varphi)$, the set $\varphi^{-1}([0, \infty))$ is a geodesic ray beginning at $P$.

The *chamber bundle* of $S$, denoted $\mathcal{C}S$, will be the set of all chambers. $G$ acts transitively on $\mathcal{C}S$ and the stabilizer of a point is conjugate to $Z_K(\mathfrak{a})$ $(= M)$. In particular, $\mathcal{C}S$ has the structure of a differentiable manifold, and it is a fiber bundle over $S$; each fiber is isomorphic to $K/Z_K(\mathfrak{a})$.

The additive group of $\mathfrak{a}$ acts in an evident way on $\mathcal{C}S$: given $X \in \mathfrak{a}$ and a chamber $(P, F, \varphi)$, one defines $X(P, F, \varphi) = (\varphi^{-1}(X), F, \varphi')$, where there is a unique choice of $\varphi'$ that makes this a continuous action. In particular, $\mathcal{C}S$ carries a natural $\mathbb{R}^r$ action. In the case $r = 1$ this is the geodesic flow on the unit tangent bundle.

Finally, if $\Gamma$ is any discrete subgroup of $G$, one sees that $\mathbb{R}^r$ acts on $\Gamma\backslash\mathcal{C}S$, which fibers over $\Gamma\backslash S$. The main result of the present paper may be phrased as follows: a measure on $\Gamma\backslash S$ arising from a limit of eigenfunction measures lifts to an $\mathbb{R}^r$-invariant measure on $\Gamma\backslash\mathcal{C}S$.

5.4. **Relation to $\Psi$DOs.** Zelditch's original proof for hyperbolic 2-space involved the construction of an equivariant pseudodifferential calculus based on the non-Euclidean Fourier transform of Helgason. It is certainly reasonable to expect that this could be generalized to higher rank; however, for the application to quantum chaos, the methods of this paper seem more efficient. In either approach, the positivity and Cartan invariance require proof.

Of course, the two methods are very closely linked. In this section we translate the representation-theoretic methods of this paper to the microlocal viewpoint. In fact, we will only do the bare minimum to show that the microlocal lifts constructed in the present paper are "compatible" with the standard construction for a general Riemannian manifold described in [2]. We will also only sketch the proof; it is more or less formal.

From the microlocal viewpoint the system under consideration resembles completely integrable systems, in that there are several commuting observable; see e.g. [20]. Of course, the Cartan flow differs from the completely integrable case in that it is very chaotic.

Initially let the notation be as in the introduction; in particular let $Y$ be a compact Riemannian manifold, $\Delta$ the Laplacian on $Y$, $S^*Y$ the unit cotangent bundle. We fix a quantization scheme $\mathrm{Op}$ that associates to a smooth function $a$ on $S^*Y$ a pseudo-differential operator $\mathrm{Op}(a)$ on $Y$ of order $0$. Let $\psi_n$ be a sequence of eigenfunctions of $\Delta$ with eigenvalues $\lambda_n \to -\infty$, s.t. the measure $\bar{\mu}_\infty = \lim_{n\to\infty} |\psi_n|^2 d\rho$ exists. Then, after possibly passing to a subsequence, the limit $a \mapsto \langle \mathrm{Op}(a)\psi_n, \psi_n\rangle$ exists for all $0$-homogeneous $a$ and defines a positive measure $\mu_\infty$ that lifts $\bar{\mu}_\infty$. We shall refer to this as a *standard microlocal lift*.

Now let us follow the notation of Section 2.1. For simplicity we shall assume $G$ simple and center-free and $\Gamma < G$ co-compact. We shall also identify $\mathfrak{g}$ and $\mathfrak{p}$ with their duals by means of the Killing form, and we will identify the tangent and cotangent bundle of $Y$ by means of the Riemannian structure (induced from the Killing form as well). We denote by $\|\cdot\|$ the norm induced on $\mathfrak{g}$ and $\mathfrak{g}^*$ by the Killing form.

Let us recall more carefully the connection between $X$ and the tangent bundle of $Y$. As before set $X = \Gamma\backslash G$, $Y = \Gamma\backslash G/K$, $S = G/K$, and let $\pi\colon X \to Y$ denote the natural projection. Let $TS$ and $TY$ denote the tangent bundles of $S$ and $Y$, and let $x_K \in S$ be the point with stabilizer $K$. Let $T^1Y \subset TY$ be the unit tangent bundle; we will often implicitly identify functions on $T^1Y$ with $0$-homogeneous functions on $TY$, and in particular functions on $T^1Y$ gives rise to pseudodifferential operators of order $0$.

We shall endow $G\times\mathfrak{p}$ with the left $G$-action given by $g(h,Y) = (gh,Y)$, and with the right $K$-action given by $(h,Y)k = (hk, k^{-1}Yk)$. There is a natural map $G \to S$ given by $g \mapsto gx_K$. This lifts to a $G$-equivariant map $G\times\mathfrak{p} \to TS$; this latter map is specified by requiring that its restriction to $\{e\}\times\mathfrak{p}$ be the usual identification of $\mathfrak{p}$ with the tangent space to $S$ at $x_K$. Taking quotients by $\Gamma$, we descend to a map also denoted $\pi\colon X\times\mathfrak{p} \to TY$. This map is constant on $K$-orbits, and factors through to a map $X\times\mathfrak{p}/K \to TY$.

In view of our identification of tangent and cotangent bundles, the symbol of a pseudodifferential operator on $Y$ may then be regarded as a $K$-invariant function on $X\times\mathfrak{p}$. We shall fix a quantization scheme $\mathrm{Op}$ that associates to such a symbol a pseudo-differential operator on $Y$.

Let $\{\psi_n\}_{n=1}^\infty \subset L^2(Y)$ be a sequence of eigenfunctions on $Y$ with parameters $\nu_n \in \mathfrak{a}^*$ and so that $\frac{\nu_n}{\|\nu_n\|} \to \tilde{\nu}$. We shall assume that $\{\psi_n\}$ is conveniently arranged in the sense of Definition 3.8. We let $\lambda_n$ be the Laplacian eigenvalue of $\psi_n$ (this differs by a constant from $-\|\nu_n\|^2$, in fact). We can and will also regard the $\psi_n$ as $K$-invariant functions on $X$. Associated to each $\psi_n$ is an $G$-intertwiner $R_n : (V_K, I_{\nu_n}) \to L^2(X)$.

We shall use $o(1)$ to denote quantities with go to 0 as $\|\nu_n\| \to \infty$.

Let other notations be as in Section 3.3. The relation between the $\Psi DO$ viewpoint and the methods of this paper are summarized in:

**Proposition.** *Let $a \in C^\infty(TY)$ be such that $a$ is 0-homogeneous. Let $g \in C^\infty(X)$ be defined by $g(x) = a(\pi(x, \tilde{\nu}))$. Suppose that $g$ is right $K$-finite. Then*

$$\langle \mathrm{Op}(a)\psi_n, \psi_n\rangle = \mu_n(\varphi_0, \delta)(g) + o(1). \tag{5.1}$$

It follows that if $\mu_{\infty,T^1Y}$ is a standard microlocal lift then $\mu_{\infty,T^1Y}$ is supported on $\pi(X \times \{\tilde{\nu}\})$, and the restriction of $\mu_{\infty,T^1Y}$ to this copy of $X$ is a microlocal lift in the sense of the current paper.

*Proof.* In three stages.

*First step.* We first verify that, if $g = 0$, then $\langle \mathrm{Op}(a)\psi_n, \psi_n\rangle = o(1)$.

Let $P$ be a $K$-invariant polynomial on $\mathfrak{p}$ of degree $d$ and consider the function $\tilde{P} : (x, A) \in X \times \mathfrak{p} \mapsto P(A)$. The function $\tilde{P}$ descends to $TY$, and there is an invariant differential operator $\mathcal{D}_P$ on $Y$ of degree $d$ whose symbol agrees with $\tilde{P}$. Since $\psi_n$ is an eigenfunction for the ring of invariant differential operators, it follows in particular that $\psi_n$ is an eigenfunction for $\mathcal{D}_P$ with eigenvalue $P(\nu_n)$. It follows that, for any $b \in C^\infty(X \times \mathfrak{p})^K$,

$$\frac{P(\nu_n)}{\|\nu_n\|^d}\langle \mathrm{Op}(b)\psi_n, \psi_n\rangle = \frac{\langle \mathrm{Op}(b)\mathcal{D}_P\psi_n, \psi_n\rangle}{\|\nu_n\|^d} = \frac{\langle \mathrm{Op}(b\tilde{P})\psi_n, \psi_n\rangle}{\|\nu_n\|^d} + o(1). \tag{5.2}$$

(5.2) implies, in particular, that if $P(\tilde{\nu}) = 0$ the statement of the Proposition holds for $a = b\tilde{P}$. We can deduce the claim of the first step by density: if $a|_{X\times\{\tilde{\nu}\}^K}$ is identically 0, then one can verify that $a$ may be densely approximated (in the topology induced by symbol-norm) by linear combinations of functions $b \cdot P$ where $P(\tilde{\nu}) = 0$. We conclude using $L^2$-bounds on pseudodifferential operators ([10, Thm. 18.1.11] and remarks after proof.)

*Second step.* We next construct an explicit class of test functions $a$ for which (5.1) holds.

Let $\sigma \in C^\infty(X)_K$, $u \in U(\mathfrak{g})$ of degree $\le d$, and let $\pi^*$ and $\pi_*$ be, respectively, the pull-back and push-forward operations on functions arising from $\pi : X \to Y$. (In other words, $\pi_*$ is obtained by integrating along $K$-orbits.) Let $\mathrm{mult}_\sigma$ be the operation "multiplication by $\sigma$" on $C^\infty(X)$. We can define by the spectral calculus of self-adjoint operators an endomorphism $(1-\Delta)^{-d/2} \colon C^\infty(Y) \to C^\infty(Y)$. We then define an endomorphism of $C^\infty(Y)$ via the rule

$$\mathrm{MyOp}(\sigma) : f \mapsto \pi_* \circ \mathrm{mult}_\sigma \circ u \circ \pi^* \circ (1-\Delta)^{-d/2} f$$

In other words, one applies $(1-\Delta)^{-d/2}$, lifts the resulting function to $X$, applies $u$ and multiplies by $\sigma$, and pushes back down to $Y$.

Regard $u$ as defining (its "symbol") a polynomial function $u_d$ of degree $d$ on $\mathfrak{g}^*$ (therefore on $\mathfrak{g}$) and let $a_{\sigma,u}$ be the following $K$-invariant function on $X \times \mathfrak{p}$:

$$a_{\sigma,u} : (x, A) \in X \times \mathfrak{p} \mapsto \frac{1}{||A||^d} \int_K \sigma(xk) u_d(k^{-1}Ak) dk$$

We verify (5.1) for $a = a_{\sigma,u}$ and $g_{\sigma,u} \stackrel{\text{def}}{=} a_{\sigma,u}|X \times \{\tilde{\nu}\}$. Note that

$$g_{\sigma,u}(x) = \int_K \sigma(xk) u_d(k^{-1}\tilde{\nu}k) dk \tag{5.3}$$

The operator $\mathrm{mult}_\sigma \circ u$ is clearly a differential operator on $X$, and one deduces that the operator $\pi^* \circ \mathrm{mult}_\sigma \circ u \circ \pi^*$ is, in fact, a differential operator on $Y$. One computes that the symbol of this latter operator is associated to the $K$-invariant function $(x, A) \in X \times \mathfrak{p} \mapsto \int_K \sigma(xk) u_d(k^{-1}Ak) dk$. We deduce that:

$$\langle \mathrm{MyOp}(\sigma)\psi_n, \psi_n\rangle - \langle \mathrm{Op}(a_{\sigma,u})\psi_n, \psi_n\rangle = o(1) \tag{5.4}$$

Further, if we regard $\psi_n$ as a $K$-invariant function on $X$:

$$\langle \mathrm{MyOp}(\sigma)\psi_n, \psi_n\rangle = (1-\lambda_n)^{-d/2} \int_X \overline{\psi_n(x)}\sigma(x)(u\psi_n)(x)\, dx \tag{5.5}$$

On the other hand, recall the definition of $\mu_n$ from Section 3.2. Let $\delta_{(N)}$ be the $N$-truncation of $\delta$ (see Definition 2.3). Choosing $N$ sufficiently large, we have $\mu_n(\varphi_0, \delta) g_{\sigma,u} = \mu_n^T(\varphi_0, \delta_N) g_{\sigma,u}$; in particular

$$\mu_n(\varphi_0, \delta) g_{\sigma,u} = \int_X dx\, \psi_n(x) \cdot \overline{R_n(\delta_N)(x)} \cdot \int_K \sigma(xk) u_d(k^{-1}\tilde{\nu}k) dk \tag{5.6}$$

$$= \int_X \int_K dx\, dk\, \psi_n(x) \overline{R_n(\delta_N)(xk^{-1})} \sigma(x) u_d(k^{-1}\tilde{\nu}k) \tag{5.7}$$

$$= \int_X dx\, \psi_n(x)\sigma(x) \overline{R_n\left(\int_K dk\, \overline{u_d(k\tilde{\nu}k^{-1})} I_{\nu_n}(k)\, \delta_N\right)} \tag{5.8}$$

At the last step, we make the substitution $k \mapsto k^{-1}$, and use the fact that the representation $I_{\nu_n}|_K$ is just the operation of right translation.

To simplify this further, we use Lemma 2.8.

Let $p_u$ be the function $k \mapsto u_d(k\tilde{\nu}k^{-1})$; it defines a function on $M\backslash K$ and thus we can regard $p_u \in V_K$. Denote by $\overline{p_u}$ the complex conjugate of $p_u$. Since $\delta_N$ is, as a function on $M\backslash K$, an approximation to a $\delta$-function, we have as $N \to \infty$: $\int_K dk\, \overline{u_d(k\tilde{\nu}k^{-1})} I_{\nu_n}(k)\delta_N \to p_u$. Here the convergence occurs in $C(M\backslash K)$. It follows that for any $n$:

$$\mu_n(\varphi_0, \delta) g_{\sigma,u} = \langle \psi_n(x)\sigma(x), R_n(\overline{p_u}) \rangle_{L^2(X)} \tag{5.9}$$

In view of the definitions, the right-hand side of (5.9) is just $\mu_n(\varphi_0, \overline{p_u})(\sigma)$. By (the proof of) Lemma 3.10, $\mu_n(\varphi_0, \overline{p_u})(\sigma) = \mu_n(p_u\varphi_0, \varphi_0)(\sigma) + o(1)$. Consequently,

$$\mu_n(\varphi_0, \delta) g_{\sigma,u} = \int_X \overline{\psi_n(x)} R_n(p_u)(x)\sigma(x)dx + o(1) \tag{5.10}$$

On the other hand, a computation with Lemma 2.8 shows that

$$p_u - \frac{I_{\nu_n}(u)\varphi_0}{(1-\lambda_n)^{d/2}} \to 0 \text{ in } L^2(M\backslash K).$$

Combining this with (5.10), we obtain:

$$\mu_n(\varphi_0, \delta) g_{\sigma,u} - (1-\lambda_n)^{-d/2} \int_X \overline{\psi_n(x)} \sigma(x) u \psi_n(x) dx = o(1) \tag{5.11}$$

In view of (5.4), (5.5) and (5.11) we have verified (5.1) in the case of $a = a_{\sigma,u}$.

*Third step.* Note that, in the statement of the Proposition, the function $g$ is necessarily right $M$-invariant. In view of what has been proved, it now suffices to check that functions of the form $g_{\sigma,u}$ (see (5.3)) span $C^\infty(X/M)_K$. This is easily reduced to checking that the linear span of the functions $k \mapsto u_d(k^{-1}\tilde{\nu}k)$ is a dense subspace of $C(M\backslash K)$. This is shown in the proof of Lemma 3.10. □

## REFERENCES

1. Jean Bourgain and Elon Lindenstrauss, *Entropy of quantum limits*, Comm. Math. Phys. **233** (2003), no. 1, 153–171.
2. Yves Colin de Verdière, *Ergodicité et fonctions propres du laplacien*, Comm. Math. Phys. **102** (1985), no. 3, 497–502. MR **87d:**58145
3. Jacques Dixmier, *Enveloping algebras*, Graduate Studies in Mathematics, vol. 11, American Mathematical Society, Providence, RI, 1996, Revised reprint of the 1977 translation. MR **97c:**17010
4. J. J. Duistermaat, J. A. C. Kolk, and V. S. Varadarajan, *Erratum: "Spectra of compact locally symmetric manifolds of negative curvature"*, Invent. Math. **54** (1979), no. 1, 101. MR **82a:**58050b
5. ______, *Spectra of compact locally symmetric manifolds of negative curvature*, Invent. Math. **52** (1979), no. 1, 27–93. MR **82a:**58050a
6. Manfred Einsiedler and Anatole Katok, *Invariant measures on $G/\Gamma$ for split simple Lie groups $G$*, Comm. Pure Appl. Math. **56** (2003), no. 8, 1184–1221, Dedicated to the memory of Jürgen K. Moser. MR **2004e:**37042

7. Sigurdur Helgason, *Groups and geometric analysis*, Mathematical Surveys and Monographs, vol. 83, American Mathematical Society, Providence, RI, 2000, Integral geometry, invariant differential operators, and spherical functions, Corrected reprint of the 1984 original. MR **2001h:**22001
8. Eberhard Hopf, *Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung*, Ber. Verh. Sächs. Akad. Wiss. Leipzig **91** (1939), 261–304. MR 1,243a
9. Lars Hörmander, *The spectral function of an elliptic operator*, Acta Math. **121** (1968), 193–218. MR 58 #29418
10. ______, *The analysis of linear partial differential operators. III*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 274, Springer-Verlag, Berlin, 1994, Pseudo-differential operators, Corrected reprint of the 1985 original. MR **95h:**35255
11. H. Iwaniec and P. Sarnak, *$L^\infty$ norms of eigenfunctions of arithmetic surfaces*, Ann. of Math. (2) **141** (1995), no. 2, 301–320. MR **96d:**11060
12. Anthony W. Knapp, *Representation theory of semisimple groups*, Princeton Mathematical Series, vol. 36, Princeton University Press, Princeton, NJ, 1986, An overview based on examples. MR **87j:**22022
13. Elon Lindenstrauss, *On quantum unique ergodicity for $\Gamma\mathbb{H} \times \mathbb{H}$*, Internat. Math. Res. Notices (2001), no. 17, 913–933. MR **2002k:**11076
14. ______, *Invariant measures and arithmetic quantum unique ergodicity*, preprint (2003), (54 pages).
15. Stephen D. Miller, *On the existence and temperedness of cusp forms for $\mathrm{SL}_3(\mathbb{Z})$*, J. Reine Angew. Math. **533** (2001), 127–169. MR **2002b:**11070
16. G. D. Mostow, *Strong rigidity of locally symmetric spaces*, Princeton University Press, Princeton, N.J., 1973, Annals of Mathematics Studies, No. 78. MR 52 #5874
17. Zeév Rudnick and Peter Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), no. 1, 195–213. MR **95m:**11052
18. A. I. Šnirel′man, *Ergodic properties of eigenfunctions*, Uspehi Mat. Nauk **29** (1974), no. 6(180), 181–182. MR 53 #6648
19. Christopher D. Sogge, *Concerning the $L^p$ norm of spectral clusters for second-order elliptic operators on compact manifolds*, J. Funct. Anal. **77** (1988), no. 1, 123–138. MR **89d:**35131
20. John A. Toth and Steve Zelditch, *$L^p$ norms of eigenfunctions in the completely integrable case*, Ann. Henri Poincaré **4** (2003), no. 2, 343–368. MR 1 985 776
21. Scott A. Wolpert, *Semiclassical limits for the hyperbolic plane*, Duke Math. J. **108** (2001), no. 3, 449–509. MR **2003b:**11051
22. Steven Zelditch, *Pseudodifferential analysis on hyperbolic surfaces*, J. Funct. Anal. **68** (1986), no. 1, 72–105. MR **87j:**58092
23. ______, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), no. 4, 919–941. MR **89d:**58129
24. ______, *The averaging method and ergodic theory for pseudo-differential operators on compact hyperbolic surfaces*, J. Funct. Anal. **82** (1989), no. 1, 38–68. MR **91e:**58194

LIOR SILBERMAN, DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NJ 08544-0001, USA.
*E-mail address*: `lior@Math.Princeton.EDU`

AKSHAY VENKATESH, DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139-4307, USA.
*E-mail address*: `akshayv@Math.MIT.EDU`