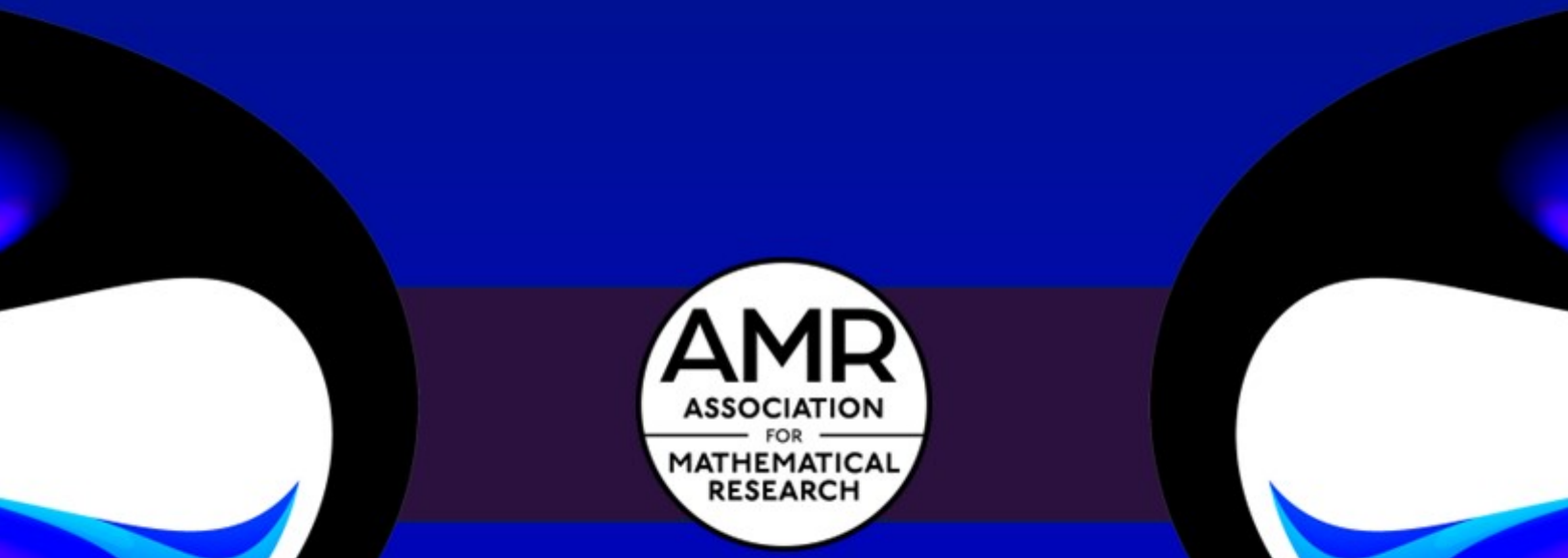
# JOURNAL OF THE ASSOCIATION FOR MATHEMATICAL RESEARCH




VOLUME 2 : ISSUE 1

2024

# JOURNAL OF THE ASSOCIATION FOR MATHEMATICAL RESEARCH



PUBLISHED BY

Association for Mathematical Research, amathr.org,
Davis, CA; Jenkintown PA.

JOURNAL OF THE ASSOCIATION FOR MATHEMATICAL RESEARCH, 2(1):15–52, 2024
*jamathr.org*

# Effective Counting in Sphere Packings

Alex Kontorovich *  Christopher Lutsko

*Received 08 Dec 2022; Revised 04 Jan 2024; Accepted 09 Feb 2024*

**Abstract:** Given a Zariski-dense, discrete group, $\Gamma$, of isometries acting on $(n+1)$-dimensional hyperbolic space, we use spectral methods to obtain a sharp asymptotic formula for the growth rate of certain $\Gamma$-orbits. In particular, this allows us to obtain a best-known effective error rate for the Apollonian and (more generally) Kleinian sphere packing counting problems, that is, counting the number of spheres in such with radius bounded by a growing parameter. Our method extends the method of Kontorovich [Kon09], which was itself an extension of the orbit counting method of Lax-Phillips [LP82], in two ways. First, we remove a compactness condition on the discrete subgroups considered via a technical cut-off and smoothing operation. Second, we develop a coordinate system which naturally corresponds to the inversive geometry underlying the sphere counting problem, and give structure theorems on the arising Casimir operator and Haar measure in these coordinates.

**Key words and phrases:** Orbital counting, Spectral theory, Automorphic representations

## 1 Introduction

The purpose of this paper is to give improved error estimates on the counting problem for Kleinian sphere packings (and discrete counting methods more broadly). A packing $\mathcal{P}$ of $\mathbb{S}^n$ (thought of as the boundary of hyperbolic $(n+1)$-space, $\mathbb{H}^{n+1}$) is an infinite collection of round, disjoint balls whose union is dense in $\mathbb{S}^n$. Following [KK23], such is called *Kleinian* if its residual set (left over when the interiors of the balls are removed) agrees with the limit set of some discrete, geometrically finite subgroup, $\Gamma$, of isometries of $\mathbb{H}^{n+1}$. A familiar example is the classical Apollonian circle packing in $n=2$, see e.g [Kon13] for more background and see Figure 1 for an example.

*Kontorovich is partially supported by NSF grant DMS-1802119, BSF grant 2020119, and the 2020-2021 Distinguished Visiting Professorship at the National Museum of Mathematics.


DOI: 10.56994/JAMR.002.001.002

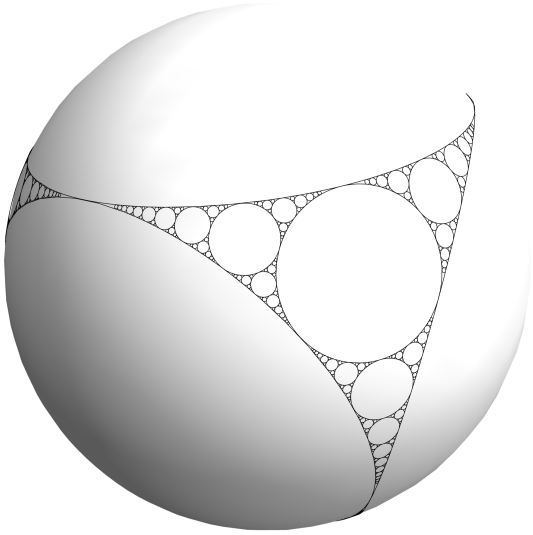


Figure 1: A classical Apollonian circle packing in the sphere $\mathbb{S}^2$. Note that the union of balls in the packing is dense in $\mathbb{S}^2$. (Image by Iván Rasskin.)

Let $\mathcal{P}$ be a given Kleinian sphere packing. For a sphere $S \in \mathcal{P}$, let $b(S)$ denote its bend, that is, its (signed) inverse-radius; this is determined after a choice of coordinates on $\mathbb{S}^n$, and in particular a choice of a point at $\infty$ (see §4). The Counting Problem is to estimate the number

$$N_{\mathcal{P}}(T) := \#\{S \in \mathcal{P} : b(S) < T\}$$

of spheres in $\mathcal{P}$ with bend bounded by a parameter $T \to \infty$. If the packing $\mathcal{P}$ is bounded, that is, the chosen point at infinity is contained in the interior of some ball, then $N_{\mathcal{P}}(T)$ is finite for all $T$; otherwise one can count spheres restricted to a bounded region (such as a period, if the packing is periodic). Let $\delta = \dim(\mathcal{P})$ be the Hausdorff dimension of the residual set of $\mathcal{P}$.

For the classical Apollonian packing, it is known that $\delta \approx 1.3$ (in general one has $n-1 < \delta < n$). In this setting, Boyd [Boy73] showed that $N_{\mathcal{P}}(T) = T^{\delta+o(1)}$, which was improved in Kontorovich-Oh [KO11] to an asymptotic formula, $N_{\mathcal{P}}(T) \sim c_{\mathcal{P}} T^{\delta}$, where $c_{\mathcal{P}}$ is a constant depending on the packing $\mathcal{P}$. An effective power savings error rate was shown in [Vin12] and [LO13] independently. These tools and results have been generalized by many authors (see, e.g., [Kim15, MO15, Pan17, EO21]). In this paper, we introduce a new method, modifying the approach in [Kon09], to produce a best known error exponent in the Counting Problem (including the classical Apollonian case).

The error exponent involves the spectrum of the hyperbolic Laplacian $\Delta$ acting on $L^2(\Gamma\backslash\mathbb{H}^{n+1})$ where $\Gamma$ is the symmetry group of the packing. From work of Lax-Phillips [LP82], Patterson [Pat76], and Sullivan [Sul84], we have that the Laplace spectrum consists of a discrete isolated bottom eigenvalue $\lambda_0 = \delta(n-\delta)$, then possibly a finite number of further discrete eigenvalues in the "exceptional" range,

$$\lambda_0 < \lambda_1 \leq \cdots \leq \lambda_k < n^2/4, \tag{1}$$

and purely continuous spectrum above $n^2/4$. (It was suggested by Sarnak [Sar07] that, in the case of the classical Apollonian packing, $k=0$, that is, there are no other discrete eigenvalues below 1 except the bottom.[1]) Write $\lambda_j = s_j(n-s_j)$ with $n/2 \le s_j < \delta$.

These eigenvalues also correspond to the parameters of spherical complementary series representations occurring in the decomposition into irreducibles of $L^2(\Gamma\backslash G)$, where $G = \mathrm{SO}(n+1,1)$, see §2.2. We assume throughout that no nonspherical complementary series representations arise in this decomposition. This holds automatically for circle packings, that is, when $n=2$. Minor modifications are needed to handle the general case, leading to potentially worse error terms, see Remark 7.

**Theorem 2.** *Given a Kleinian packing $\mathcal{P}$ as above, there exists a constant $c_{\mathcal{P}} > 0$ such that:*

$$N_{\mathcal{P}}(T) = c_{\mathcal{P}} T^{\delta}\left(1 + O\left(T^{-\eta}\right)\right) \tag{3}$$

*as $T \to \infty$, where*

$$\eta = \frac{2}{n+3}(\delta - s_1). \tag{4}$$

*If there is no discrete spectrum other than the base eigenvalue, then we have that*

$$N_{\mathcal{P}}(T) = c_{\mathcal{P}} T^{\delta}\left(1 + O\left(T^{-\eta}(\log T)^{2/(n+3)}\right)\right),$$

*where $\eta = \frac{2}{n+3}(\delta - n/2)$.*

For the classical Apollonian packing ($n=2$), our error exponent is $\eta = \frac{2}{5}(\delta - s_1)$, and if there are no discrete eigenvalues above the base then $\eta = \frac{2}{5}(\delta - 1) \approx 0.12\ldots$, whereas the best previously known exponent [LO13, Theorem 1.1] was:

$$\eta = \frac{2}{63}(\delta - s_1), \tag{5}$$

and if $s_1 = n/2$, this exponent is $\eta = \frac{2}{63}(\delta - 1) \approx 0.0097$. Hence (4) is a significant improvement over (5).

**Remark 6.** *Counting with a smooth cutoff and extracting all of the lower order terms corresponding to eigenvalues other than the base, we obtain the best possible error exponent $\eta = \delta - n/2$ (see Theorem 61), which improves over the smooth error exponent in the Apollonian case $\eta = \frac{2}{7}(\delta - 1)$ in [LO13, Theorem 8.2].*

**Remark 7.** *If we remove the assumption that there are no nonspherical complementary series representations occurring in $L^2(\Gamma\backslash G)$, then our method gives the weaker error term in (3) with $\eta = \frac{2}{n+3}(\delta - (n-1))$, that is, replacing $s_1$ in (4) with $n-1$; see Remark 79. With more effort, this error term could be improved, see [EO21].*

The proof of Theorem 2 introduces two new technical ideas: first, of independent interest, we overcome difficulties in the "abstract spectral theory" method of [Kon09] arising from a non-compactness condition (see the discussion below), and second, we introduce a new decomposition tailored to sphere counting problems and derive structure theorems for the arising Casimir operator and Haar measure.

[1] Added in print: Sarnak's question has now been answered in the affirmative [KKL23].

## 1.1 Ideas in the Proof

**Previous Approach:** To explain the main ideas, we first recall the method introduced in [Kon09], which itself is modeled on [LP82]. Consider the following related but simpler to exposit problem. Let $\Gamma < \mathrm{SL}_2(\mathbb{R})$ be a discrete, Zariski dense, geometrically finite subgroup, with critical exponent $\delta_\Gamma > 1/2$, acting on the upper half plane $\mathbb{H}^2$, and assume that $\infty$ is not a *point of approximation* for $\Gamma$. This last condition implies that either $\infty$ is a cusp of $\Gamma$ with stabilizer $\Gamma_\infty$ (in which case the limit set is periodic), or $\infty$ is not in the limit set of $\Gamma$ (and hence the limit set is bounded). Either way, consider the problem of counting

$$N_\Gamma(T) := \#\{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_\infty\backslash\Gamma : c^2+d^2 < T\}.$$

(The Lax-Phillips work counts $a^2+b^2+c^2+d^2 < T$, and new ideas are needed to handle the potential stabilizer when counting $c^2+d^2$.) The assumption that $\infty$ is not a point of approximation assures that the set of such $c^2+d^2$ is discrete, and hence this count is finite for any $T$. The main idea, as sketched below, is to use a particular function of the Laplacian to grow balls of size $T$ from balls of bounded size.

In this subsection, write $G = \mathrm{SL}_2(\mathbb{R})$, $N = \{\left(\begin{smallmatrix} 1 & \mathbb{R} \\ 0 & 1 \end{smallmatrix}\right)\}$, (so that $\Gamma_\infty = \Gamma \cap N$) and $K = \mathrm{SO}(2)$, and let

$$\chi_T : G \to \mathbb{R} : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \mathbb{1}_{\{c^2+d^2<T\}} \tag{8}$$

be the indicator function of the region in question in $G$. Note that $\chi_T$ is left-$N$-invariant and right-$K$-invariant, and let

$$F_T : \Gamma\backslash G/K \to \mathbb{R} : g \mapsto \sum_{\gamma \in \Gamma_\infty\backslash\Gamma} \chi_T(\gamma g),$$

so that $N_\Gamma(T) = F_T(e)$. As observed in [Kon09], $F_T$ is in $L^2(\Gamma\backslash G/K) = L^2(\Gamma\backslash\mathbb{H})$ if and only if $\infty$ is a cusp of $\Gamma$. To access the value of $F_T$ at the origin, we let $\psi$ be a smooth bump function about the origin in $G$, and automorphize it to $\Psi(g) := \sum_{\gamma\in\Gamma} \psi(\gamma g)$. Then, we can write a smooth approximation of $N_\Gamma(T)$ as

$$\widetilde{N_\Gamma}(T) := \langle F_T, \Psi\rangle_\Gamma$$

where the inner-product is with respect to $L^2(\Gamma\backslash\mathbb{H})$. Now suppose we take the inner-product of $F_T$ with an eigenfunction of the Laplacian $\phi$, with eigenvalue $\lambda = s(1-s)$. Then, solving a second order ODE, we would have that

$$\langle F_T, \phi\rangle_\Gamma = \alpha T^s + \beta T^{1-s}, \tag{9}$$

for some $\alpha$ and $\beta$ depending on $\phi$. The key idea of [Kon09] is then to rewrite (9) in a way that involves the eigenvalue only, and not the coefficients $\alpha,\ \beta$ of the eigenfunction. This is achieved by setting $T = 1$ and $T = b$ (for some $b < 10$, see §3.1) in (9) and solving for $\alpha, \beta$, to give an expression of the form:

$$\langle F_T, \phi\rangle = K_T(s)\langle F_1, \phi\rangle + L_T(s)\langle F_b, \phi\rangle,$$

for some functions $K_T, L_T$. This then allows one to prove the "main identity", which states that, in the sense of $L^2$, we have:

$$F_T = K_T(\Delta)F_1 + L_T(\Delta)F_b.$$

This is exactly what we mean by "growing" the ball of radius $T$ from the Laplacian and bounded norm balls. It is this identity that can be proved rigorously even in the absence of explicit Whittaker expansions and spectral decompositions.

**New Ideas:** Much of this approach fails if $\Gamma$ does not have a cusp at $\infty$, and the main purpose of [KO12] was to bypass this "PDE" approach and replace it with homogeneous dynamics, at the cost of worse error terms. The main new ideas of this paper allow us to recover the PDE approach (and then extend it to the Kleinian setting).

The first issue is that $F_T$ is not in $L^2$. The observation which eliminates this issue (made already in [KO12]) is to add a second cut off in the $N$-direction without losing any of the orbit, since the limit set is anyway bounded; this adds compactness in the $x$-variable, restricted to some sufficiently large interval $[-X,X]$. Now we again unfold the inner product $\langle F_T, \Psi\rangle_\Gamma$ leading to the following integral:

$$\int_0^\infty \int_{-X}^{X} \chi_T(z)\widetilde{\chi}_X(z)\Psi(z)dx\frac{dy}{y^2}.$$

We proceed with harmonic analysis on $[-X,X]\times(0,\infty)$. However this truncation requires a delicate smoothing procedure to avoid introducing boundary terms in the analysis of the Laplacian. Once this is accomplished, the proof follows in a similar fashion. See §3.1 for the details.

In addition to overcoming the restriction in [Kon09], this $\mathrm{SL}_2(\mathbb{R})$ result is of independent interest, and improves on [KO12, Theorem 1.8] which used methods from homogeneous dynamics to count Pythagorean triples.

**In the Kleinian Setting:** There are several further modifications and innovations needed to extend the above-described $\mathrm{SL}_2(\mathbb{R})$ approach to the setting of orbits in circle/sphere packings. In the previous setting, the stabilizer of $\chi_T$ in (8) was left-$N$ and right-$K$ invariant, and so it was natural to use Iwasawa coordinates $\mathrm{SL}_2(\mathbb{R}) = NAK$. In the setting of sphere packings, one counts spheres in the orbit $S_0\Gamma$ with bend less than $T$; here $S_0$ is a fixed $(n-1)$-sphere in $\widehat{\mathbb{R}^n} = \mathbb{R}^n \cup \{\infty\} = \partial\mathbb{H}^{n+1}$ and $\Gamma < G = \mathrm{SO}(n+1,1)$ is a symmetry group of the packing, acting on the right by Möbius transformations. The analogous function $\chi_T$ is given by:

$$\chi_T : G \to \mathbb{R} : g \mapsto \mathbb{1}_{\{b(S_0g)<T\}},$$

where again $b(S)$ is the bend of a sphere $S$. This function is left-$H$ invariant, where $H = \mathrm{Stab}_G(S_0) \cong \mathrm{SO}(n,1)$; it is also right invariant under the group $L$ of rigid affine motions, since neither translating nor rotating a sphere changes its bend. The latter decomposes further as $L = UM$, where $U$ is a one-parameter unipotent group (which controls the co-bend, defined to be the bend of the inversion of a sphere through the unit sphere), and $M \cong \mathrm{SO}(n)$ rotates the sphere about the origin. It turns out that $H \cap M \cong \mathrm{SO}(n-1)$, and thus we set $M_1 := M/(H\cap M) \cong \mathbb{S}^{n-1}$. The subgroup of $G$ which directly controls the bend is also a one-parameter unipotent group we call $\overline{U}$, leading to the map:

$$H \times \overline{U} \times U \times M_1 \to G,$$

which is an isomorphism in a neighborhood of the identity; see §4.2 for details. An important feature of this decomposition is the fact that the Haar measure of $G$ in these coordinates decomposes as a product of $H$-Haar measure on the $H$ component, times the $M_1$-Haar measure on the $M_1$ component, and times something depending only on $U$ and $\overline{U}$; see §4.4. Moreover, in the proof, we only need the Casimir

operator restricted to left-$H$- and right-$M_1$-invariant functions, for which we derive a nice, concise explicit expression in any dimension; see §4.5.

Let $\Gamma_1 := \Gamma \cap H$ denote the stabilizer of $S_0$ in $\Gamma$. It follows from the Structure Theorem for Kleinian packings [KK23, Theorem 22] that $\Gamma_1$ is a *lattice* in $H$ (that is, it acts with finite covolume); this fact will be used crucially in our analysis. The finiteness of the volume of $\Gamma_1\backslash H$ is analogous in the $\mathrm{SL}_2(\mathbb{R})$ setting to the finiteness of the volume of $K$, though the latter is trivial since $K$ is compact. Note that we acted on the left in $\mathrm{SL}_2(\mathbb{R})$ and it is more convenient to act on the right for sphere packings.

In the $\mathrm{SL}_2(\mathbb{R})$ setting, the $N$ direction was unbounded and required a cut-off. Analogously here, the $U$-direction is unbounded; this can be controlled via a similar truncation procedure by invoking the fact that the limit set is bounded in the $U$ direction, see Lemma 64.

**Remark 10.** *Note that if the stabilizer of $\mathcal{P}$ contains a full rank unipotent subgroup then the methods of [Kon09] may be applied directly. For the existence of such, see [KN19].*

### 1.2 Organization

In section 2 we collect some preliminaries. In section 3, we warm up to the counting theorem and illustrate the main ideas by proving a result analogous to Theorem 2 in the $\mathrm{SL}_2(\mathbb{R})$ setting. In section 4, we switch to the general $\mathrm{SO}(n+1,1)$ setting, and derive the Haar measure, and Casimir operator in the above-described coordinate system. Finally, in section 5 we prove Theorem 2.

## 2 Preliminaries

### 2.1 Lie algebras and the Casimir Operator

We collect here some standard facts about the group $G = SO^\circ(n+1,1)$ (where $\circ$ denotes the connected component of the identity), and its Lie algebra $\mathfrak{g} = Lie(G)$ of dimension $d := \dim(\mathfrak{g}) = (n+1)(n+2)/2$.

In general, Casimir operators generate the center of the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$. In our rank-one setting, we can compute the Casimir operator $\mathcal{C}$ as follows. Let $X_1,\dots,X_d$ be a basis for the Lie algebra $\mathfrak{g}$, and let $X_1^*,\dots,X_d^*$ be a dual basis with respect to the Killing form:

$$B(X,Y) = \mathrm{Tr}(\mathrm{Ad}(X)\circ \mathrm{Ad}(Y)),$$

that is, $B(X_i,X_j^*) = \mathbb{1}_{\{i=j\}}$. Then the Casimir operator can be expressed as

$$\mathcal{C} = \sum_{i=1}^{d} X_i X_i^*.$$

Since the elements of the Lie algebra act like first order differential operators, the Casimir operator acts as a second order differential operator on smooth functions on $G$; see §4.5 for a detailed calculation in our setting.

Let $K \cong SO(n+1) < G$ be a maximal compact subgroup. When restricted to $K$-invariant smooth functions on $G$, the Casimir operator $\mathcal{C}$ agrees (up to constant) with the hyperbolic Laplacian $\Delta$ under an identification $G/K \cong \mathbb{H}^{n+1}$.

## 2.2 Decomposition of $L^2(\Gamma\backslash G)$ into irreducibles

Let $\Gamma$ be a discrete, geometrically finite, Zariski dense subgroup of $G$. Given Iwasawa coordinates $G = NAK$, let $M \cong \mathrm{SO}(n)$ be the centralizer of $A$ in $K$. The group $G$ acts by the right-regular representation on the Hilbert space $\mathcal{H} := L^2(\Gamma\backslash G)$ of square-integrable $\Gamma$-automorphic functions. Recalling the standing assumption that $\mathcal{H}$ does not weakly contain any nonspherical complementary series representations, the space $\mathcal{H}$ decomposes into components as follows:

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_k \oplus \mathcal{H}^{tempered}. \tag{11}$$

Here each $\mathcal{H}_j$ is the $G$-span of the eigenfunction corresponding to the eigenvalue $\lambda_j = s_j(n-s_j)$ in (1), each of which is an irreducible spherical complementary series representation with parameter $s_j > n/2$, and $\mathcal{H}^{tempered}$ denotes tempered spectrum. Moreover, the subspace $\mathcal{H}_j^K$ of $K$-fixed vectors in $\mathcal{H}_j$ is 1-dimensional, and spanned by the corresponding eigenfunction in $L^2(\Gamma\backslash\mathbb{H}^{n+1}) \cong L^2(\Gamma\backslash G)^K$. In general, nonspherical complementary series representations can only occur if $n \geq 3$ and parameter $s \leq n-1$ [Kna01].

## 2.3 Abstract Spectral Theorem

We recall the abstract spectral theorem (see for example [Rud73, Ch. 13]) for unbounded self-adjoint operators. Let $L$ be a self-adjoint, positive semidefinite operator on the Hilbert space $\mathcal{H}$. In our applications $\mathcal{H}$ will be either $L^2(\Gamma\backslash G)$ or a subspace thereof and $L$ will be the Casimir operator.

**Theorem 12** (Abstract Spectral Theorem)**.** *There exists a spectral measure $\nu$ on $\mathbb{R}$ and a unitary spectral operator $\widehat{\ }: \mathcal{H} \to L^2([0,\infty), d\nu)$ such that:*

*[label = )]Abstract Parseval's Identity: for $\phi_1, \phi_2 \in \mathcal{H}$*

$$\langle \phi_1, \phi_2 \rangle_{\mathcal{H}} = \langle \widehat{\phi_1}, \widehat{\phi_2} \rangle_{L^2([0,\infty), d\nu)}; \tag{13}$$

*The spectral operator is diagonal with respect to L: for $\phi \in \mathcal{H}$ and $\lambda \geq 0$,*

$$\widehat{L\phi}(\lambda) = \lambda \widehat{\phi}(\lambda); \tag{14}$$

Moreover, if $\lambda$ is in the point specturm of $L$ with associated eigenspace $\mathcal{H}_\lambda$, then for any $\psi_1, \psi_2 \in \mathcal{H}$ one has

$$\widehat{\psi_1}(\lambda)\widehat{\psi_2}(\lambda) = \langle \mathrm{Proj}_{\mathcal{H}_\lambda} \psi_1, \mathrm{Proj}_{\mathcal{H}_\lambda} \psi_2 \rangle, \tag{15}$$

where Proj refers to the projection to the subspace $\mathcal{H}_\lambda$. In the special case that $\mathcal{H}_\lambda$ is one-dimensional and spanned by the normalized eigenfunction $\phi_\lambda$, we have that

$$\widehat{\psi_1}(\lambda)\widehat{\psi_2}(\lambda) = \langle \psi_1, \phi_\lambda \rangle \langle \phi_\lambda, \psi_2 \rangle. \tag{16}$$

## 3 The $\mathrm{SL}_2(\mathbb{R})$ Case

In this section, let $G := \mathrm{SL}_2(\mathbb{R})$ and $\Gamma < G$ be a Zariski dense, finitely generated, discrete subgroup with $\delta_\Gamma > 1/2$. The goal of this section is to prove Theorem 18 below, that is, to improve on [KO12, Theorem 1.11] by extending the proof of [Kon09, Theorem 1.3 (1)] to the setting where $\Gamma_\infty$ is trivial. This will serve as a model for the method that we will generalize to higher dimensions in the rest of the paper.

Again, from work of Lax-Phillips [LP82] and Patterson [Pat76] we have that the Laplace spectrum below $1/4$ consists of a finite number of discrete eigenvalues

$$\lambda_0 = \delta_\Gamma(1-\delta_\Gamma) < \lambda_1 \leq \cdots \leq \lambda_k < 1/4. \tag{17}$$

Write $\lambda_j = s_j(1-s_j)$ with $s_j \in (\frac{1}{2}, 1)$.

**Theorem 18.** *Let $G := \mathrm{SL}_2(\mathbb{R})$ and $\Gamma < G$ a Zariski dense, finitely generated, discrete subgroup with $\delta_\Gamma > 1/2$. Assume that $\infty$ is either a cusp for $\Gamma$ with stabilizer $\Gamma_\infty$ or $\infty$ is not in the limit set of $\Gamma$. Then there exist constants $c_0 > 0$, $c_1, \ldots, c_k$, and $\eta > 0$ such that as $T \to \infty$*

$$\begin{aligned} N_\Gamma(T) &:= \#\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_\infty \backslash \Gamma : c^2 + d^2 < T \right\} \\ &= c_0 T^{\delta_\Gamma} + c_1 T^{s_1} + \cdots + c_k T^{s_k} + O(T^\eta \log^{1/2} T), \end{aligned} \tag{19}$$

*with $\eta = \frac{1}{2}(\delta_\Gamma + \frac{1}{2})$. (Of course some of the lower order "main terms", if any, may be dominated by the error term.)*

**Remark 20.** *The case where $\infty$ is a cusp is the content of [Kon09, Theorem 1.3 (1.5)], so we assume below that $\infty$ is not in the limit set of $\Gamma$.*

**Remark 21.** *Note that in this case, since we are counting points in $\mathbb{H}^2$ (which is $K$ invariant) we can extract all lower order terms corresponding to eigenvalues other than the base.*

**Remark 22.** *The corresponding error term in [KO12, §4.1] is significantly worse (and not even explicitly specified) as compared to Theorem 18, due in part to much worse dependence on Sobolev norms of the corresponding test vectors.*

To begin the proof, we proceed as described in the Introduction. For $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$, let

$$\chi_T(g) := \begin{cases} 1 & \text{if } c^2 + d^2 < T, \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

Under the identification $G/K \sim \mathbb{H}$ (where $K = SO(2)$), $g \mapsto z = gi$, $\chi_T$ corresponds to the condition $\Im(z) > 1/T$.

Now average $\chi_T$ over the group $\Gamma$:

$$F_T(g) := \sum_{\gamma \in \Gamma} \chi_T(\gamma g). \tag{24}$$

such that the count in (19) can be written as $N_\Gamma(T) = F_T(e)$.

To access the value of $F_T$ at the identity, we follow the standard procedure of smoothing the count. To this end, we fix once and for all a smooth, even, nonnegative, compactly supported bump function $\psi_1 \in C_0^\infty(\mathbb{R})$ with unit total mass, $\int_{\mathbb{R}} \psi_1 dx = 1$. Given $\varepsilon > 0$, we set $\psi_\varepsilon(x) := \frac{1}{\varepsilon}\psi_1(\frac{x}{\varepsilon})$. Write Iwasawa coordinates as $n_x := \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right) \in N$, $a_y := \mathrm{Diag}(\sqrt{y}, 1/\sqrt{y}) \in A$, and $k \in K = \mathrm{SO}(2)$, and by abuse of notation, write $\psi : G \to \mathbb{R}_+$ as follows:

$$\psi(n_x a_y k) := \psi_\varepsilon(x)\psi_\varepsilon(\log y). \tag{25}$$

Clearly $\psi$ is right-$K$-invariant and depends on $\varepsilon$ (although we omit this from the notation). It is easy to compute that $\int_G \psi(n_x a_y k)\frac{dx\,dy\,dk}{y^2} = 1 + O(\varepsilon)$.

We automorphise $\psi$ by setting:

$$\Psi(g) = \Psi_\varepsilon(g) := \sum_{\gamma\in\Gamma} \psi(\gamma g),$$

and consider the smoothed count:

$$\widetilde{N}_\Gamma(T) = \widetilde{N}_\Gamma^\varepsilon(T) := \langle F_T, \Psi\rangle.$$

After unfolding $F_T$, we see that

$$\widetilde{N}_\Gamma(T) := \sum_{\gamma\in\Gamma} w_T(\gamma),$$

where $w_T = w_{T,\varepsilon} : \mathbb{H} \to [0,1]$ is given by

$$w_T(g) := \int_G \chi_T(gh)\psi(h)dh. \tag{26}$$

The following theorem, from which Theorem 18 follows by optimizing error terms, gives an asymptotic expansion for the smoothed count in the $\mathrm{SL}_2(\mathbb{R})$-setting.

**Theorem 27.** *Assume that $\varepsilon > 0$ is small enough. Then there exist constants $c_{\Gamma,\varepsilon}^{(i)}$ for $i = 0, 1, \ldots, k$ such that*

$$\widetilde{N}_\Gamma(T) = c_{\Gamma,\varepsilon}^{(0)} T^{\delta_\Gamma} + c_{\Gamma,\varepsilon}^{(1)} T^{s_1} + \cdots + c_{\Gamma,\varepsilon}^{(k)} T^{s_k} + O\left(\frac{1}{\varepsilon} T^{1/2}\log T\right) \tag{28}$$

*with $c_{\Gamma,\varepsilon}^{(0)} > 0$, and the implied constant depending only on $\Gamma$. Moreover $c_{\Gamma,\varepsilon}^{(i)} = c^{(i)}(1 + O(\varepsilon))$ for all $i = 0, 1, \ldots, k$.*

It remains to prove Theorem 27.

### 3.1 Inserting the Laplacian

The smooth count above is an $L^2$ inner product of the indicator function $F_T$ with a smooth bump function $\Psi$. The key idea now is to forget about $\Psi$ and analyze the structure of the inner product of $F_T$ with *any* smooth $L^2$ function using the hyperbolic Laplacian $\Delta = -y^2(\partial_{xx}+\partial_{yy})$.

Following [Kon09, §3] let

$$K_T(s) := \frac{T^s b^{1-s} - T^{1-s} b^s}{b^{1-s} - b^s}, \qquad L_T(s) := \frac{T^{1-s} - T^s}{b^{1-s} - b^s}, \tag{29}$$

where $1 < b < 3$ is a constant depending on $T$ which ensures that $K_T$ and $L_T$ have no poles (see [Kon09, (3.7)] for an exact computation). With that, if $\phi$ *were* an eigenfunction of the Laplacian then we would have $\langle F_T, \phi\rangle = A_\phi T^s + B_\phi T^{1-s}$. In this case, we could conclude that

$$\langle F_T, \phi\rangle = K_T(s)\langle F_1, \phi\rangle + L_T(s)\langle F_b, \phi\rangle.$$

Thus we would like to show that

$$F_T = K_T(\Delta)F_1 + L_T(\Delta)F_b, \tag{30}$$

where $K_T(\Delta)$ (and $L_T(\Delta)$) is defined via (29) in the same way that one defines matrix exponentiation (via Taylor expansion). However, there is a problem created by the fact that $F_T$ is not in $L^2(\Gamma\backslash\mathbb{H})$. To get around this, we will first restrict the support of $F_T$ in the $x$-direction before applying various smoothing arguments to conclude that a version of (30) holds for the modified $F_T$. Since $\infty$ is not in the limit set, it is in the free boundary, and hence there exists a fixed an $X = X(\Gamma) > 0$ such that the full region $((-\infty,-X]\cup[X,\infty)) \times [0,\infty) \subset \mathbb{H}$ is contained in a single fundamental domain, see Figure 2.

**Restricting the Real Direction:** Define the following counting function

$$F_{T,X}(z) := \sum_{\gamma\in\Gamma} \chi_T(\gamma z)\widetilde{\chi}_X(\gamma z)$$

where

$$\widetilde{\chi}_X(x+iy) := \begin{cases} 1 & \text{if } x \in [-X,X], \\ 0 & \text{if } x \notin [-X,X]. \end{cases}$$

Let $\mathcal{I}_X := [-X,X]$ denote the support of $\widetilde{\chi}_X$. Note that by our choice of $X$, we still have that $N_\Gamma(T) = F_T(i) = F_{T,X}(i)$.

Now, consider the difference operator

$$G_{T,X} := F_{T,X} - K_T(\Delta)F_{1,X} - L_T(\Delta)F_{b,X}.$$

Our goal is to show, that for any $\Psi \in L^2(\Gamma\backslash\mathbb{H})$ we have $\langle G_{T,X}, \Psi\rangle_\Gamma = 0$. This implies the following identity

**Proposition 31.** *For any values of $T$ and $X$ large enough depending on the group $\Gamma$, we have*

$$F_{T,X} = K_T(\Delta)F_{1,X} + L_T(\Delta)F_{b,X}. \tag{32}$$

*where $K_T$ and $L_T$ are the differential operators defined above and $b$ fixed.*

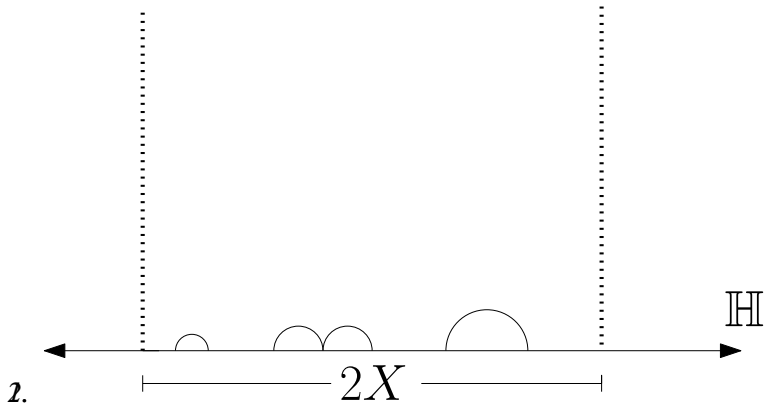



Figure 2: A fundamental domain of $\Gamma$ which extends to infinity in the real direction, and the cut-off due to $\widetilde{\chi}_X$ which is large enough such that $[-X,X]$ contains the entire limit set.

Proving this proposition will be somewhat involved, so we break it down into several steps. First, we unfold the inner product

$$\langle G_{T,X}, \Psi \rangle = \int_{\mathbb{H}} (\chi_T(z)\widetilde{\chi}_X(z) - K_T(\Delta)\chi_1(z)\widetilde{\chi}_X(z) - L_T(\Delta)\chi_b(z)\widetilde{\chi}_X(z))\Psi(z)\frac{dxdy}{y^2}.$$

**Smoothing $\chi_T$:** Now we smooth the function $\chi_T$ so that we can move the Laplacian over to $\widetilde{\chi}_X$ via self-adjointness. Let $\sigma > 0$, and define the following smooth cut-off function

$$\chi_{1,\sigma}(x+iy) := \begin{cases} 1 & \text{if } y > 1, \\ 0 & \text{if } y < (1-\sigma), \end{cases}$$

and smooth, in between the two bounds. Now let $\chi_{T,\sigma}(x+iy) := \chi_{1,\sigma}(x+iTy)$.

Let $G^{\sigma}_{T,X}$ be defined similarly to $G_{T,X}$, with $\chi_T$ replaced by $\chi_{T,\sigma}$. That is, let

$$G^{\sigma}_{T,X}(\Psi) := \int_0^{\infty} \int_{x\in\mathcal{I}_X} \widetilde{\chi}_X(z) \\ (\chi_{T,\sigma}(z)\Psi(z) - \chi_{1,\sigma}(z)(K_T(\Delta)\Psi)(z) - \chi_{b,\sigma}(z)(L_T(\Delta)\Psi)(z))\, dx\frac{dy}{y^2}.$$

Note that by construction, for any fixed $t$,

$$\int_0^{\infty} \int_{x\in\mathcal{I}_X} |\chi_{t,\sigma}(z) - \chi_t|^2\, dx\frac{dy}{y^2} \ll_{t,X} \sigma.$$

Thus, using Cauchy-Schwarz, we have that for any $\Psi \in L^2(\Gamma\backslash\mathbb{H})$

$$\lim_{\sigma\to 0} G^{\sigma}_{T,X}(\Psi) = \langle G_{T,X}, \Psi \rangle_{\Gamma}.$$

Note also that, by the support of the $\chi_{*,\sigma}$, we may also write, when convenient,

$$G^{\sigma}_{T,X}(\Psi) := \int_{1/2T}^{\infty}\int_{x\in\mathcal{I}_X} \widetilde{\chi}_X(z) \\ (\chi_{T,\sigma}(z)\Psi(z) - \chi_{1,\sigma}(z)(K_T(\Delta)\Psi)(z) - \chi_{b,\sigma}(z)(L_T(\Delta)\Psi)(z))\,dx\frac{dy}{y^2}.$$

Now our goal, is to establish the following lemma, note that we have already fixed $T$, $X$ large enough, $\sigma > 0$ small enough.

**Lemma 33.** *For any function $\Psi \in L^2(\Gamma\backslash\mathbb{H})$ we have that $G^{\sigma}_{T,X}(\Psi) = 0$.*

To prove this lemma, fix $\sigma$, then we will show that for any $\varepsilon > 0$ small enough $\left|G^{\sigma}_{T,X}(\Psi)\right| < \varepsilon$.

**Periodizing and Smoothing $\Psi$:** Now we periodize and smooth $\Psi$ in order to do spectral theory on $\mathcal{I}_X \times (0,\infty)$ which we call the space $\mathcal{F}_X$. Note that $\mathcal{F}_X \cong \Xi_\infty\backslash\mathbb{H}$ for the elementary group $\Xi_\infty := \left(\begin{smallmatrix}1 & 2X\mathbb{Z}\\ 0 & 1\end{smallmatrix}\right)$, but this coordinate description is more useful when we generalize to higher dimensions. Let $L^2(\mathcal{F}_X)$ denote the $L^2$ space with measure $\frac{dxdy}{y^2}$ and periodic with respect to $\Xi_\infty$. Let $\widetilde{\Psi}(z)$ denote a function which agrees with $\Psi$ on

$$[-X, X-\eta)\times(1/2T,\infty) \subset \mathbb{H},$$

where $\eta > 0$ is to be chosen later. From $x = X-\eta$ to $x = X-\eta/2$, $\widetilde{\Psi}(z)$ smoothly interpolates between the values of $\Psi(z)$ and $\Psi(z-2X)$, and from $x = X-\eta/2$ up to $x = X$, $\widetilde{\Psi}$ is exactly equal to $\Psi(z-2X)$, ensuring that all derivatives of $\widetilde{\Psi}$ at the boundary values $x = X$ and $x = -X$ agree. We also impose that $\widetilde{\Psi}$ decays rapidly to zero below $y = 1/4T$, so that $\widetilde{\Psi} \in L^2(\mathcal{F}_X)$.

Note that the cost of moving from $\Psi$ to $\widetilde{\Psi}$ is small. Indeed,

$$\begin{aligned}\mathrm{diff}(\Psi) &:= \int_{1/2T}^{\infty}\int_{X-\eta}^{X}\left|\widetilde{\Psi}(z)-\Psi(z)\right|^2 dx\frac{dy}{y^2}\\ &\ll \int_{1/2T}^{\infty}\int_{X-\eta}^{X}|\Psi(z-2X)|^2+|\Psi(z)|^2 dx\frac{dy}{y^2}.\end{aligned}$$

Since $\Psi \in L^2(\Gamma\backslash\mathbb{H})$, as $\eta\to 0$, this integral goes to zero, and hence by choosing $\eta$ small enough, we can make the difference less than $\varepsilon^2$.

Now note that

$$\begin{aligned}&\int_{1/2T}^{\infty}\int_{\mathcal{I}_X}(\chi_{T,\sigma}(z)\widetilde{\chi}_X(z) - K_T(\Delta)\widetilde{\chi}_X(z)\chi_{1,\sigma}(z) - L_T(\Delta)\widetilde{\chi}_X(z)\chi_{b,\sigma}(z))\Psi(z)dx\frac{dy}{y^2}\\ &= \int_{1/2T}^{\infty}\int_{\mathcal{I}_X}(\chi_{T,\sigma}(z)\widetilde{\chi}_X(z) - K_T(\Delta)\chi_{1,\sigma}(z)\widetilde{\chi}_X(z) - L_T(\Delta)\chi_{b,\sigma}(z)\widetilde{\chi}_X(z))(\Psi(z)-\widetilde{\Psi}(z))dx\frac{dy}{y^2}\\ &+ \int_{1/2T}^{\infty}\int_{\mathcal{I}_X}(\chi_{T,\sigma}(z)\widetilde{\chi}_X(z) - K_T(\Delta)\chi_{1,\sigma}(z)\widetilde{\chi}_X(z) - L_T(\Delta)\chi_{b,\sigma}(z)\widetilde{\chi}_X(z))\widetilde{\Psi}(z)dx\frac{dy}{y^2},\end{aligned}$$

and by Cauchy Schwarz and the above argument

$$\int_{1/2T}^{\infty}\int_{\mathcal{I}_X}(\chi_{T,\sigma}(z)\widetilde{\chi}_X(z)-K_T(\Delta)\chi_{1,\sigma}(z)\widetilde{\chi}_X(z)-L_T(\Delta)\chi_{b,\sigma}(z)\widetilde{\chi}_X(z))(\Psi(z)-\widetilde{\Psi}(z))dx\frac{dy}{y^2}$$
$$\ll(\mathrm{diff}(\Psi))^{1/2}\left[\int_{1/2T}^{\infty}\int_{\mathcal{I}_X}\Big(|(\chi_{T,\sigma}(z)\widetilde{\chi}_X(z)|^2+|K_T(\Delta)\chi_{1,\sigma}(z)\widetilde{\chi}_X(z)|^2 + |L_T(\Delta)\chi_{b,\sigma}(z)\widetilde{\chi}_X(z))|^2\Big)dx\frac{dy}{y^2}\right]^{1/2}$$

Since $\mathrm{diff}(\Psi)$ is bounded by $\varepsilon^2$, it remains to show that the integral in the brackets is finite. Taking, say, the middle term (the others being similar), we note that, since the region $[1/2T,\infty)\times\mathcal{I}_X$ is contained in a finite number (depending on $T$) of translates of a fixed fundamental domain, $\mathcal{F}$, for $\Gamma$,

$$\int_{1/2T}^{\infty}\int_{\mathcal{I}_X}|K_T(\Delta)\chi_{1,\sigma}(z)\widetilde{\chi}_X(z)|^2dx\frac{dy}{y^2}\le\sum_{\gamma\in\Gamma}\int_{\mathcal{F}}|K_T(\Delta)\chi_{1,\sigma}(\gamma z)\widetilde{\chi}_X(\gamma z)|^2\frac{dxdy}{y^2},$$

where only finitely many values of $\gamma$ give nonzero contributions to the sum. For each contributing $\gamma$, we can extend the function $h_\gamma(z)=h_{\gamma,T,X,\sigma}(z):=\chi_{1,\sigma}(\gamma z)\widetilde{\chi}_X(\gamma z)$ supported on $\mathcal{F}$ to a $\Gamma$-automorphic function $H_\gamma(z)$ on $\mathbb{H}$, which agrees with $h_\gamma(z)$ on $\mathcal{F}$. Applying the Abstract Spectral Theorem (13), (14) in $L^2(\Gamma\backslash\mathbb{H})$, we get

$$\int_{\Gamma\backslash\mathbb{H}}|K_T(\Delta)H_\gamma(z)|^2\frac{dxdy}{y^2}=\int_{\lambda\in[0,\infty)}\left|K_T(\lambda)\widehat{H_\gamma}(\lambda)\right|^2d\nu(\lambda)$$
$$\ll T^2\|H_\gamma\|^2_{L^2(\Gamma\backslash\mathbb{H})}\ll_{T,X,\sigma}1,$$

where $\nu$ is the abstract spectral measure, and we used the bounds $K_T(\lambda)\ll T^s\le T^1$.

With that, the difference

$$\left|G^\sigma_{T,X}(\Psi)-G^\sigma_{T,X}(\widetilde{\Psi})\right|\ll_{T,X,\sigma}\varepsilon \tag{34}$$

for any value of $\varepsilon>0$. Thus it remains to show that $G^\sigma_{T,X}(\widetilde{\Psi})=0$. Since $\widetilde{\Psi}(x+iy)$ is periodic on $x\in\mathcal{I}_X$, we can use self-adjointness to write

$$G^\sigma_{T,X}(\widetilde{\Psi})=\int_0^{\infty}\int_{\mathcal{I}_X}(\chi_{T,\sigma}(z)-K_T(\Delta)\chi_{1,\sigma}(z)-L_T(\Delta)\chi_{b,\sigma}(z))\widetilde{\Psi}(z)dx\frac{dy}{y^2}. \tag{35}$$

**Working on $\mathcal{F}_X$:** To prove (35) let Now, to prove that $G^\sigma_{T,X}(\widetilde{\Psi})=0$, let

$$g^\sigma_{T,X}(z):=\chi_{T,\sigma}(z)-K_T(\Delta)\chi_{1,\sigma}(z)-L_T(\Delta)\chi_{b,\sigma}(z).$$

Thus, we have

$$G^\sigma_{T,X}(\widetilde{\Psi})=\langle g^\sigma_{T,X},\widetilde{\Psi}\rangle_{\mathcal{F}_X}.$$

Lemma 33 will then follow from the following lemma, which shows that, for an arbitrary $\psi\in L^2(\mathcal{F}_X)$ the inner product $\langle g^\sigma_{T,X},\psi\rangle$ is not correlated with any almost eigenfunction:

**Lemma 36.** *Fix $T$, $X$, and $\sigma$ as above, then for any $\psi \in L^2(\mathcal{F}_X)$ and any $\lambda \geq 0$, we have*

$$\langle g^{\sigma}_{T,X}, \psi \rangle_{\mathcal{F}_X} \ll_{\lambda,T,\sigma,X} \|(\Delta - \lambda)\psi\|_{\mathcal{F}_X}. \tag{37}$$

*Proof of Lemma 36.* Fix $\psi \in L^2(\mathcal{F}_X)$ and consider

$$\langle \chi_{T,\sigma}, \psi \rangle_{\mathcal{F}_X} = \int_0^{\infty} \chi_{T,\sigma}(y) \int_{\mathcal{I}_X} \psi(z) dx \frac{dy}{y^2},$$

and define $f(y) := \int_{\mathcal{I}_X} \psi(z)dx$ and $h(y) := \int_{\mathcal{I}_X} (\Delta - \lambda)\psi(z)dx$. Then we note that, by periodicity in the $x$-direction,

$$\begin{aligned} h(y) &= \int_{\mathcal{I}_X} (\Delta - \lambda)\psi(z)dx \\ &= -y^2 \partial_{yy} f(y) - \lambda f(y). \end{aligned}$$

Thus, [Kon09, Lemma B.1] (which is a simple application of the method of variation of parameters) implies

$$f(y) = \alpha y^s + \beta y^{1-s} + y^s u(y) + y^{1-s} v(y), \tag{38}$$

where

$$u(y) := (1-2s)^{-1} \int_{(1-\sigma)/T}^{y} w^{-1-s} h(w) dw,$$

$$\text{and} \quad v(y) := (1-2s)^{-1} \int_{(1-\sigma)/T}^{y} w^{s-2} h(w) dw,$$

if $\lambda = s(1-s) \neq 1/4$. Note that the choice of integration boundary for $u$ and $v$ corresponds to the bottom of the support of $\chi_{T,\sigma}$; this will be convenient later on. Moreover when $\lambda = 1/4$

$$f(y) = \alpha y^{1/2} + \beta y^{1/2} \log y + y^{1/2} u(y) + v(y) y^{1/2} \log y, \tag{39}$$

where

$$u(y) := \int_{(1-\sigma)/T}^{y} w^{-3/2} \log(w) h(w) dw,$$

$$\text{and} \qquad v(y) := -\int_{(1-\sigma)/T}^{y} w^{-3/2} h(w) dw.$$

Therefore (assuming $\lambda \neq 1/4$ for simplicity) integrating the $y$ variable gives

$$\langle \chi_{T,\sigma}, \psi \rangle_{\Xi_\infty} = \int_0^{\infty} \chi_{T,\sigma}(y)(\alpha y^s + \beta y^{1-s}) \frac{dy}{y^2} + I + II$$

where $I := \int_0^\infty \chi_{T,\sigma}(y) y^{s-2} u(y) dy$ and $II := \int_0^\infty \chi_{T,\sigma}(y) y^{-1-s} v(y) dy$. Now we can use Cauchy-Schwarz (as in [Kon09, (B.5)]) to establish that $I, II \ll \|(\Delta - \lambda)\psi\|_{\mathcal{F}_X}$. In fact, this is the crucial reason why we needed to work on $\mathcal{F}_X$ which is compact in the $x$ direction, and thus $\chi_{T,\sigma}$ is (square) integrable. Thus, we may conclude

$$\langle \chi_{T,\sigma}, \psi \rangle_{\mathcal{F}_X} = A_\sigma T^s + B_\sigma T^{1-s} + O(\|(\Delta - \lambda)\psi\|_{\mathcal{F}_X})$$

where $A_\sigma = \beta \int_0^\infty \chi_{1,\sigma}(y) y^{-1-s} dy$ and $B_\sigma := \alpha \int_0^\infty \chi_{1,\sigma}(y) y^{s-2} dy$.

Given our choice of $K_T$ and $L_T$ from (29) we have that

$$K_T(s), L_T(s) \ll \begin{cases} T^s & \text{if } s \in (1/2, 1], \\ T^{1/2} \log T & \text{if } s = 1/2 + it. \end{cases} \tag{40}$$

Then, by the analysis in [Kon09, Proposition 3.5] we have

$$\langle g^\sigma_{T,X}, \psi \rangle_{\mathcal{F}_X} \ll_{T,\lambda,X,\sigma} \|(\Delta - \lambda)\psi\|_{\mathcal{F}_X} \tag{41}$$

for any choice of $\psi \in L^2(\mathcal{F}_X)$. □

From there we can choose $\psi$ to be as in [Kon09, Proof of Theorem 3.2] to establish that $g^\sigma_{T,X}$ is almost everywhere 0. Note that for this argument to work one only needs the bound (41), a Hilbert space (here $L^2(\mathcal{F}_X)$), an unbounded self-adjoint operator (i.e $\Delta$), and the abstract spectral theorem.

From there we conclude that

$$\langle g^\sigma_{T,X}, \widetilde{\Psi} \rangle_{\mathcal{F}_X} = 0$$

From there, we conclude that $G^\sigma_{T,X}(\widetilde{\Psi}) = 0$. Then thanks to (34) we conclude

$$\begin{aligned} G^\sigma_{T,X}(\Psi) &= \int_0^\infty \int_{\mathcal{I}_X} \widetilde{\chi}_X(z)(\chi_{T,\sigma}(z) - K_T(\Delta)\chi_{1,\sigma}(z) - L_{T,\sigma}(\Delta)\chi_{b,\sigma}(z))\Psi(z) dx \frac{dy}{y^2} \\ &\ll_{\sigma,T,X} \varepsilon \end{aligned}$$

for any value of $\varepsilon > 0$. Taking $\varepsilon$ to 0 establishes Lemma 33. Since we have that for any $\Psi \in L^2(\Gamma\backslash\mathbb{H})$

$$\lim_{\sigma \to 0} G^\sigma_{T,X}(\Psi) = \langle G_{T,X}, \Psi \rangle_\Gamma$$

we conclude that $\langle G_{T,X}, \Psi \rangle_\Gamma = 0$. Which is exactly Proposition 31. □

## 3.2 Proof of Theorem 27

For the proof of Theorem 27, we return to the smoothed count

$$\widetilde{N}_\Gamma(T) = \langle F_{T,X}, \Psi \rangle_\Gamma.$$

Now apply the abstract Parseval's identity (13)

$$\begin{aligned}\langle F_{T,X},\Psi\rangle_\Gamma &= \langle \widehat{F_{T,X}},\widehat{\Psi}\rangle_{\mathrm{Spec}(\Gamma)}\\ &= \widehat{F_{T,X}}(\lambda_0)\widehat{\Psi}(\lambda_0)+\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}\widehat{F_{T,X}}(\lambda)\widehat{\Psi}(\lambda)\mathrm{d}\nu(\lambda),\end{aligned}\tag{42}$$

where for ease of exposition, we assume that $\lambda_0$ is the only eigenvalue below $1/4$; in general, the other eigenvalues are dealt with similarly.

Addressing the first term in (42), we can use the abstract spectral theorem, and the fact that $F_{T,X}$ and $\Psi$ are $K$-fixed, to say

$$\widehat{F_{T,X}}(\lambda_0)\widehat{\Psi}(\lambda_0)=\langle F_{T,X},\phi_0\rangle\langle\Psi,\phi_0\rangle,$$

where $\phi_0$ is the base eigenfunction. Note that, for the first factor we can apply the main identity Proposition 31, and the definition to conclude

$$\widehat{F_{T,X}}(\lambda_0)\widehat{\Psi}(\lambda_0)=T^\delta c\langle\phi_0,\Psi\rangle_\Gamma+O(T^{1/2})$$

for some constant $c$ independent of $T$. As for the second factor, by the mean value theorem (see [Kon09, (4.17)] for details) we have

$$\langle\phi_0,\Psi_\varepsilon\rangle_\Gamma=\phi_0(i)+O(\varepsilon).$$

It remains to bound the contribution to (42) from the remainder of the spectrum (assuming here that there are no isolated eigenvalues apart from the base). Using Proposition 31 we have that

$$\begin{aligned}\mathrm{Err}&:=\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}\widehat{F}_{T,X}(\lambda)\widehat{\Psi}(\lambda)d\nu\\ &=\int_{\mathrm{Spec}(\Gamma\setminus\{\lambda_0\}}\left(K_T\widehat{(\Delta)F_{1,X}}(\lambda)+L_T\widehat{(\Delta)F_{b,X}}(\lambda)\right)\widehat{\Psi}(\lambda)d\nu.\end{aligned}$$

By the abstract spectral theorem and (40) we have $K_T\widehat{(\Delta)F_{1,X}}(\lambda)\ll T^{1/2}\log T\widehat{F_{1,X}}(\lambda)$. Thus, by Cauchy-Schwarz, positivity, and Parseval's identity

$$\begin{aligned}&\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}K_T(\lambda)\widehat{F_{1,X}}(\lambda)\widehat{\Psi}(\lambda)\mathrm{d}\nu\ll T^{1/2}\log T\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}\widehat{F_{1,X}}(\lambda)\widehat{\Psi}(\lambda)\mathrm{d}\nu\\ &\ll T^{1/2}\log T\left(\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}\left|\widehat{F_{1,X}}(\lambda)\right|^2\mathrm{d}\nu\right)^{1/2}\left(\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}}\left|\widehat{\Psi}(\lambda)\right|^2\mathrm{d}\nu\right)^{1/2}\\ &\ll T^{1/2}\log T\|F_{1,X}\|_\Gamma\|\Psi\|_\Gamma,\end{aligned}$$

Since $\Psi$ is an $\varepsilon$-approximation to the identity, and since the term involving $L_T$ can be treated similarly, we thus conclude

$$\mathrm{Err}\ll\frac{X^{1/2}}{\varepsilon}T^{1/2}\log T,$$

note that $X$ does not go to $\infty$, it is kept fixed. This completes the proof of Theorem 27.

# 4 Kleinian Sphere Packings

Turning now to the sphere packing setting, let $\mathcal{P}$ be a fixed bounded Kleinian sphere packing. Given a sphere $S \in \mathcal{P}$ let $b(S)$ denote the bend of $S$. Then our aim is to establish the asymptotic, (5) for

$$N_{\mathcal{P}}(T) := \#\{S \in \mathcal{P} : b(S) < T\}.$$

## 4.1 Preliminaries on Inverse Coordinate Systems and the Symmetry Group

We now give a model of hyperbolic space, and develop an inversive coordinate system for the spheres on the ideal boundary of such; see, e.g., [KK23, §3.1] for background. To begin, we fix a real quadratic form $Q$ of signature $(n+1,1)$. For concreteness, we can change variables over $\mathbb{R}$ to the "standard" example of $Q = -x_0x_{n+1} + x_1^2 + \cdots + x_n^2$ which has half-Hessian

$$Q = \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & I_n & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}. \tag{43}$$

Then the quadratic space $(V = \mathbb{R}^{n+1,1}, Q)$ with product $v \star w = v \cdot Q \cdot w^t$ contains the cone $V_0 := \{v : Q(v) = 0\}$, the one-sheeted hyperboloid $V_1 = \{v : Q(v) = 1\}$, and the two-sheeted hyperboloid $V_{-1} := \{v : Q(v) = -1\}$; fix either component of the latter for our model of $\mathbb{H}^{n+1}$. The group $O_Q(\mathbb{R})$ acts on $V_{-1}$, and its subgroup $G = O_Q^\circ(\mathbb{R})$, that is, the connected component of the identity, fixes the components.

There is a 1-1 correspondence between vectors $v \in V_1$ and (oriented) spheres in $\partial\mathbb{H}^{n+1}$, obtained as follows. Given such a $v$, the orthogonal space $v^\perp := \{w \in V : v \star w = 0\}$ intersects the fixed component of $V_{-1}$ at a hyperplane $\cong \mathbb{H}^n$, and the ideal boundary of the latter is the desired sphere.

This geometric correspondence is made algebraic after a choice of (inversive) coordinates on $V_1$ as follows. Let $V^* := \{v^* : V \to \mathbb{R}, \text{ linear}\}$ be the dual space to $V$, and $Q^*$ the dual form, so that $v^* \star w^* = v \star w$. Fix a non-zero null covector $b^* \in V^*$, that is, $Q^*(b^*) = 0$. Also fix a null covector $\hat{b}^*$, with $b^* \star \hat{b}^* = -2$. For the case of the standard form in (43), one can make the choice $b^* = (0,\ldots,0,-2)$ and $\hat{b}^* = (-2,0,\ldots,0)$. Then one picks an orthonormal system, $bx_1^*,\ldots,bx_n^*$, for the orthogonal complement to the span of $b^*$ and $\hat{b}^*$. Then the sphere corresponding to the vector $v \in V_1$ has bend (that is, inversive radius)

$$b^*(v), \tag{44}$$

and center $\frac{1}{b^*(v)}(bx_1^*(v),\ldots,bx_n^*(v))$. When $b^*(v) = 0$, the sphere is a plane (which has no "center"), so the expression $(bx_1^*(v),\ldots,bx_n^*(v))$ is a unit normal to the plane. The "co-bend" of a sphere is defined as the bend of the image of the sphere on inversion through the unit sphere at the origin. The co-bend of the sphere corresponding to $v$ is given by $\hat{b}^*(v)$. Therefore the tuple $(b^*(v), bx_1^*(v),\ldots,bx_n^*(v),\hat{b}^*(v))$ gives a complete inversive coordinate system on $V_1$. It is sometimes convenient to isolate the 'bend-center', comprising the coordinates not including the first and last; so we define:

$$bz^*(v) := (bx_1^*(v),\ldots,bx_n^*(v)). \tag{45}$$

A Kleinian sphere packing decomposes into finitely many $\Gamma$-orbits (by the Structure Theorem [KK23, Theorem 22]), which in the above coordinate system corresponds simply to orbits $v_0 \cdot \Gamma$, with $v_0 \in V_1$,

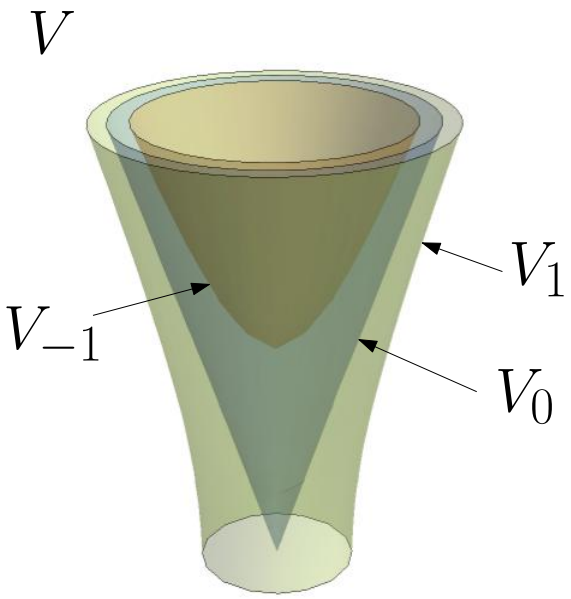



Figure 3: The quadratic space $V$, and (the upper parts of) its components $V_0$, $V_1$ and $V_{-1}$.

and the bends of such are measured by $b^*(v)$, for $v \in v_0 \cdot \Gamma$. We count the whole packing by counting one orbit at a time.

### 4.2 Decomposition of $G$

Fix $v_0$ as above. Let $\chi_T(g) = \mathbb{1}\{b^*(v_0 g) < T\}$ denote the indicator function of the vector $v_0 g$ having bend at most $T$, where $g \in G$. This function is left-invariant under $H := \mathrm{Stab}_G(v_0) = \{g \in G : v_0 g = v_0\}$, and also right-invariant under $L := \mathrm{Stab}_{b^*} = \{g \in G : b^*(vg) = b^*(v), \text{ for all } v \in V\}$.

It will be useful to decompose $G = O_Q^\circ(\mathbb{R})$ (for $Q$ given by (43)) and its Lie algebra $\mathfrak{g} := Lie(G)$ as follows. First we decompose $\mathfrak{g}$ as:

$$\mathfrak{g} = \mathfrak{h} \oplus \bar{\mathfrak{u}} \oplus \mathfrak{u} \oplus \mathfrak{m}_1.$$

Here $\mathfrak{h} = Lie(H)$, and with

$$M := \begin{pmatrix} 1 & 0 & 0 \\ 0 & O(n) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

we have that $M \cap H \cong O(n-1)$. Then set $\mathfrak{m} := Lie(M)$ and $\mathfrak{m}_1 := \mathfrak{m} \ominus Lie(M \cap H)$. (Note that $\mathfrak{h} \cap \mathfrak{m}$) is

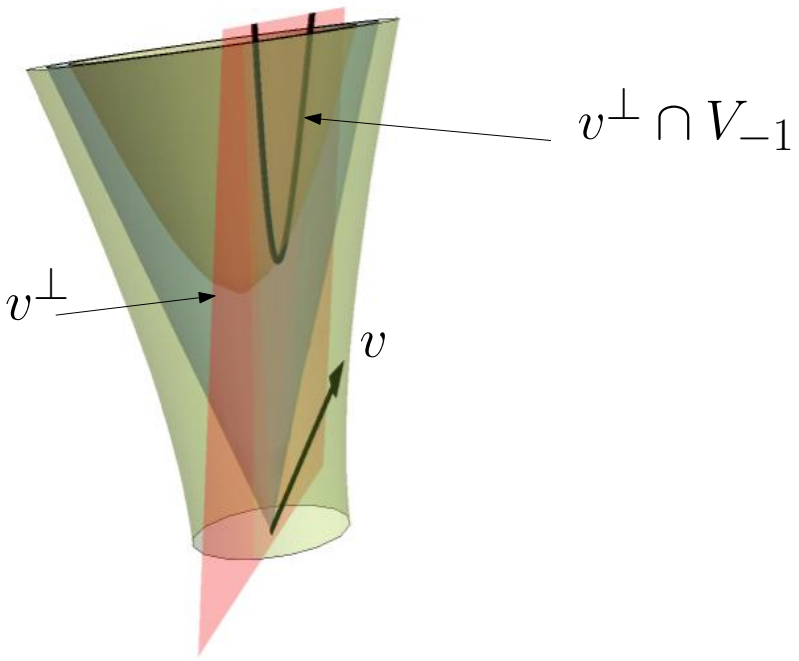



Figure 4: A vector $v \in V_1$, its orthogonal space $v^\perp$ and the intersection this orthogonal space with $V_{-1}$ which corresponds to an ideal sphere.

trivial when $G \cong O(3,1)$.) Set $M_1 := \exp \mathfrak{m}_1 < G$. The one-parameter Lie algebras $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are given by:

$$\mathfrak{u} = \left\{ \begin{pmatrix} 0 & 0_{n-1} & w & 0 \\ 0 & 0 & 0 & 0_{n-1} \\ 0 & 0 & 0 & 2w \\ 0 & 0 & 0 & 0 \end{pmatrix} : w \in \mathbb{R} \right\},$$

and

$$\bar{\mathfrak{u}} = \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0_{n-1} & 0 & 0 & 0 \\ y & 0 & 0 & 0 \\ 0 & 0_{n-1} & y/2 & 0 \end{pmatrix} : y \in \mathbb{R} \right\},$$

which exponentiate to the groups

$$U := \left\{ u(w) := \begin{pmatrix} 1 & 0_{n-1} & w & w^2 \\ 0 & I_{n-1} & 0 & 0_{n-1} \\ 0 & 0 & 1 & 2w \\ 0 & 0 & 0 & 1 \end{pmatrix} : w \in \mathbb{R} \right\},$$

and

$$\overline{U} := \left\{ \overline{u}(y) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0_{n-1} & I_{n-1} & 0 & 0 \\ y & 0 & 1 & 0 \\ y^2/4 & 0_{n-1} & y/2 & 1 \end{pmatrix} : y \in \mathbb{R} \right\},$$

respectively.

This gives the corresponding decomposition:

$$H \times \overline{U} \times U \times M_1 \to G : (h, \overline{u}, u, m_1) \mapsto h\overline{u}um_1.$$

Geometrically this decomposition is in fact very natural. $H$ is the stabilizer of $v_0$, corresponding to the sphere whose orbit we are considering. The action of $\overline{U}$ changes the bend of the sphere, while $UM_1$ changes the co-bend by moving the center via a polar coordinate description of the plane ($U$ as the radial coordinate and $M_1 \cong O(n)/O(n-1)$ giving a rotation).

Let $d := \dim(\mathfrak{g}) = (n+1)(n+2)/2$. Let $p := \dim(\mathfrak{h}) = n(n+1)/2$, and let $\ell := \dim(\mathfrak{m}_1) = n-1$. Hence $d = p + \ell + 2$.

### 4.3 An Explicit Basis.

Before proceeding, it will help our calculations to fix a basis for $\mathfrak{h}$ and $\mathfrak{m}_1$. To ease notation, we will describe a basis in a special case when $n = 4$, so $d = 15$, $p = 10$ and $\ell = 3$; the general case will be clear by analogy. Take:

$$\left.\begin{matrix} X_1 := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&-1&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, X_2 := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&0&0 \\ 0&-1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, X_3 := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&1&0&0&0 \\ 0&-1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, \end{matrix}\right\} \in \mathfrak{h} \cap \mathfrak{m}$$

$$\left.\begin{matrix} X_4 := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&\frac{1}{2}&0&0 \end{pmatrix}, X_5 := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&\frac{1}{2}&0&0&0 \end{pmatrix}, X_6 := \begin{pmatrix} 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&\frac{1}{2}&0&0&0&0 \end{pmatrix}, \end{matrix}\right\} \text{Lower Triangular in } \mathfrak{h}$$

$$\left.\begin{matrix} X_7 := \begin{pmatrix} 0&0&0&1&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&2 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, X_8 := \begin{pmatrix} 0&0&1&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&2 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, X_9 := \begin{pmatrix} 0&1&0&0&0&0 \\ 0&0&0&0&0&2 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \end{pmatrix}, \end{matrix}\right\} \text{Upper Triangular in } \mathfrak{h}$$

$$X_{10} := \begin{pmatrix} 1&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&-1 \end{pmatrix}, \text{ – Diagonal in } \mathfrak{h}$$

$$X_{11} := \begin{pmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&0&0&0&\frac{1}{2}&0 \end{pmatrix} \in \overline{\mathfrak{u}}, \; X_{12} := \begin{pmatrix} 0&0&0&0&1&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&2 \\ 0&0&0&0&0&0 \end{pmatrix} \in \mathfrak{u},$$

$$X_{13} := \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&0&-1&0&0\\ 0&0&0&0&0&0 \end{pmatrix}, X_{14} := \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&0\\ 0&0&-1&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix}, X_{15} := \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&-1&0&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix} \Bigg\} \quad \in \mathfrak{m}_1.$$

To move back to group elements we denote

$$\begin{aligned} &h_j(x_j) := \exp(x_j X_j) \in H && \text{for } j \le p,\\ &\overline{u}(y) := \exp(yX_{p+1}) \in \overline{U},\\ &u(w) := \exp(wX_{p+2}) \in U,\\ &m_j(\varphi_j) := \exp(\varphi_j X_{j+p+2}) \in M_1 && \text{for } 1 \le j \le \ell. \end{aligned}$$

It will also be convenient to record the following components of $H$:

$$H_M := \prod_{X_j \in \mathfrak{h}\cap\mathfrak{m}} \exp(x_j X_j),\ H_+ := \prod_{X_j \text{ "Upper Triangular"}} \exp(x_j X_j), \tag{46}$$

$$H_- := \prod_{X_j \text{ "Lower Triangular"}} \exp(x_j X_j), \text{ and } H_A := \prod_{X_j \text{ "Diagonal"}} \exp(x_j X_j).$$

## 4.4 Calculating the Haar Measure

Given that decomposition of the Lie algebra, we now derive an explicit form of the Haar measure in our chosen coordinate system. Let

$$\mathbf{z} := (x_1, \dots, x_p, y, w, \varphi_1, \dots, \varphi_\ell)$$

denote a set of coordinates and let

$$\mathcal{J} : \mathbb{R}^d \to G : \mathbf{z} \mapsto h_1(x_1)h_2(x_2)\dots h_p(x_p)\overline{u}(y)u(w)m_1(\varphi_1)\dots m_\ell(\varphi_\ell)$$

map our coordinate space to $G$. Then, in these new coordinates we have the following decomposition of the Haar measure

**Theorem 47.** *[Haar Measure Structure Theorem] Let $\rho(\mathbf{z})$ denote the density of the Haar measure in $\mathbf{z}$ coordinates. Then*

$$\rho(\mathbf{z}) = \rho_H(x_1, \dots x_p)\rho_{\overline{U}U}(y,w)\rho_{M_1}(\varphi_1, \dots, \varphi_\ell), \tag{48}$$

*where $\rho_H$ and $\rho_{M_1}$ are (respectively) the densities for the Haar measure on $H$ and $M_1$. Moreover we have that $\rho_{\overline{U}U}(y,w) = |1+wy|^{n-1}$.*

*Proof.* Given a linear differential operator $T_{\mathbf{z}}$ acting on functions of $G$, we can represent it as

$$T_{\mathbf{z}} = \eta_1\partial_{x_1} + \dots + \eta_p\partial_{x_p} + \eta_{p+1}\partial_y + \eta_{p+2}\partial_w + \eta_{p+3}\partial_{\varphi_1} + \dots + \eta_d\partial_{\varphi_\ell}.$$

Consider an element in the Lie algebra $\mathfrak{g}$

$$Z = b_1X_1 + b_2X_2 + \dots + b_dX_d.$$

Then $Z$ corresponds to a linear differential operator via the exponential map:

$$T_{\mathbf{z}}(f(g)) = \frac{d}{dt} f(ge^{Zt})\big|_{t=0}.$$

Hence we have two representations of the linear differential operator corresponding to $Z$ (one in terms of the $b$ variables and one in terms of the $\eta$ variables), our goal is to recover the $\eta$ variables from the $b$ variables.

To first order, we can write the differential $T_{\mathbf{z}}$ acting on a function depending on $g = h_1 \cdots h_p \bar{n} n m_1 \cdots m_\ell$ as

$$\begin{aligned}
&h_1(I+\eta_1 X_1)\cdots h_p(I+\eta_p X_p)\bar{u}(I+\eta_{p+1}X_{p+1}) \\
&\qquad \cdot u(I+\eta_{p+2}X_{p+2})m_1(I+\eta_{p+3})\cdots m_\ell(I+\eta_d) \\
&= h_1\cdots h_p\bar{n}nm_1\cdots m_\ell\left\{I+\eta_1\operatorname{Ad}((h_2\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_1\right. \\
&\quad \left.+\eta_2\operatorname{Ad}((h_3\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_2+\cdots+\eta_{d-1}\operatorname{Ad}(m_\ell^{-1})X_{d-1}+\eta_d X_d\right\}.
\end{aligned}$$

Giving us the differential operator

$$\begin{aligned}
&D(\mathcal{J}) : \eta_1\partial_{x_1}+\cdots+\eta_p\partial_{x_p}+\eta_{p+1}\partial_y+\eta_{p+2}\partial_w+\eta_{p+3}\partial_{\varphi_1}+\cdots+\eta_d\partial_{\varphi_\ell} \\
&\quad \mapsto \eta_1\operatorname{Ad}((h_2\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_1+ \\
&\eta_2\operatorname{Ad}((h_3\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_2+\cdots+\eta_{p+\ell+1}\operatorname{Ad}(m_\ell^{-1})X_{d-1}+\eta_d X_d.
\end{aligned}$$

Thus, if we want to apply the differential operator $X_j$ to a function $f(g)$, then we simply solve for $\eta_i$ on the right hand side of this map. Then the left hand side tells us the action on the coordinates. Let us denote the vector of such $\eta_i$'s by $(\eta_{j1},\eta_{j2},\ldots,\eta_{jd})$. Thus

$$\begin{aligned}
X_j &= \eta_{j1}\operatorname{Ad}((h_2\cdots h_p\bar{n}nm_1\cdots m_\ell)^{-1})X_1+\eta_{j2}\operatorname{Ad}((h_3\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_2 \\
&\quad +\cdots+\eta_{j(d-1)}\operatorname{Ad}(m_\ell^{-1})X_{d-1}+\eta_{jd}X_d,
\end{aligned}$$

for every $j=1,\ldots,d$.

Now to calculate the Haar measure, we proceed by the following methodology: define the right multiplication operator $R_A(\mathbf{z}) := \mathcal{J}^{-1}(\mathcal{J}(\mathbf{z})\cdot A)$ and its Jacobian:

$$\left[R_A'(\mathbf{z})\right]_{jp} := \frac{\partial^j R_A(\mathbf{z})}{d\mathbf{z}_p}.$$

Then our goal is to find $\rho(\mathbf{z})$ such that:

$$\int f(\mathcal{J}(\mathbf{z}))\rho(\mathbf{z})d\mathbf{z} = \int f(\mathcal{J}(R_A(\mathbf{z})))\rho(\mathbf{z})d\mathbf{z}. \tag{49}$$

Changing variables $\mathbf{y} = R_{A^{-1}}(\mathbf{z})$, on the left hand side gives:

$$\int f(\mathcal{J}(\mathbf{z}))\rho(\mathbf{z})d\mathbf{z} = \int f(\mathcal{J}(R_A(\mathbf{y})))\rho(R_A(\mathbf{y}))\left|\det R_A'(\mathbf{y})\right|d\mathbf{y}$$

which is equal to the right hand side of (49). Hence we want to find $\rho$ such that:

$$\rho(R_A(\mathbf{y})) = \frac{\rho(\mathbf{y})}{\left|\det R'_A(\mathbf{y})\right|}$$

for all choices of $\mathbf{y}$ and $A \in G$. In particular, we may choose $\mathbf{y} = 0$, and $A = \mathcal{J}(\mathbf{z})$, which gives

$$\rho(\mathbf{z}) = \frac{1}{\left|\det R'_{\mathcal{J}(\mathbf{z})}(0)\right|}. \tag{50}$$

Note that since the Haar measure is only unique up to a constant, we set $\rho(0) = 1$ without loss of generality. Using our decomposition of the Lie algebra we can write:

$$[R'_{\mathcal{J}(\mathbf{z})}(0)]_{ij} = \left[\frac{\partial}{\partial_{t_i}} \mathcal{J}(\mathbf{z}) e^{\sum_{i=1}^{d} t_i X_i}\Big|_{\mathbf{t}=0}\right]_j$$

Now we can linearize the exponential and then find the corresponding coordinate description of the differential operator as we did above. The above argument implies

$$\rho(\mathbf{z})^{-1} = \det\left|[R'_{\mathcal{J}(\mathbf{z})}(0)]_{ij}\right| = \det|\eta_{ij}|.$$

Now to simplify matters, we express each adjoint as a linear combination of elements in the basis:

$$\begin{aligned}
\mathrm{Ad}((h_2 \cdots h_p \bar{u} u m_1 \cdots m_\ell)^{-1}) X_1 &= \sum_{i=1}^{d} \mu_{1i} X_i \\
\mathrm{Ad}((h_3 \cdots h_p \bar{u} u m_1 \cdots m_\ell)^{-1}) X_2 &= \sum_{i=1}^{d} \mu_{2i} X_i \\
&\cdots \\
X_d &= \sum_{i=1}^{d} \mu_{di} X_i.
\end{aligned}$$

Thus, to calculate the Haar measure, by linearity of the adjoint operator, we need to find a $d \times d$ matrix $\eta$ such that

$$\eta \mu \mathbf{X} = \mathbf{X}$$

where $\mathbf{X} := (X_1, X_2, \ldots X_d)^T$. Hence, since the determinant is multiplicative, we have that $\rho(\mathbf{z}) = \det[\mu]$.

Now suppose we wanted to calculate the Haar measure of $H$, then we can write

$$\begin{aligned}
\mathrm{Ad}((h_2\cdots h_p)^{-1})X_1 &= \sum_{i=1}^{d} \nu_{1i}X_i \\
\mathrm{Ad}((h_3\cdots h_p)^{-1})X_2 &= \sum_{i=1}^{d} \nu_{2i}X_i \\
&\cdots \\
\mathrm{Ad}(h_p^{-1})X_{p-1} &= \sum_{i=1}^{d} \nu_{(p-1)i}X_i \\
X_p &= \sum_{i=1}^{d} \nu_{pi}X_i.
\end{aligned}$$

For completeness write $\nu_{ji} = \delta_{i,j}$ for $j > p$. Then the above argument implies that one can write the Haar measure on $H$ as $\rho_H(\mathbf{z}) = \det[\nu]$.

Now the key observation is to write

$$\begin{aligned}
\mathrm{Ad}((\bar{u}um_1\cdots m_\ell)^{-1})X_1 &= \sum_{i=1}^{d} c_{1i}X_i \\
&\cdots \\
\mathrm{Ad}((\bar{u}um_1\cdots m_\ell)^{-1})X_p &= \sum_{i=1}^{d} c_{pi}X_i \\
\mathrm{Ad}((um_1\cdots m_\ell)^{-1})X_{p+1} &= \sum_{i=1}^{d} c_{(p+1)i}X_i \\
&\cdots \\
X_d &= \sum_{i=1}^{d} c_{di}X_i.
\end{aligned}$$

By linearity of the adjoint operator we have that $\mu = \nu c$, and hence $\det(\mu) = \det(\nu)\det(c)$. Or rather, $\rho(\mathbf{z}) = \rho_H(x_1,\ldots,x_p)f(w,y,\varphi_1,\ldots,\varphi_\ell)$ for some function $f$.

Moreover, since the Haar measure is unimodular, and since $M_1$ is a group, we can apply the same argument to $M_1$ on the right, and show that $\rho$ satisfies the product structure from (48).

Finally, thanks to this product structure, we can take $m_1,\ldots,m_l$ equal to the identity when calculating

$\rho_{\overline{U}U}$. Thus, let

$$\begin{aligned} \mathrm{Ad}((\overline{u}u)^{-1})X_j &= \sum_{i=1}^{d} d_{ji}X_i, \qquad \text{for } j=1,\dots,p \\ \mathrm{Ad}((u)^{-1})X_{p+1} &= \sum_{i=1}^{d} d_{(p+1)i}X_i \\ X_j &= \sum_{i=1}^{d} d_{ji}X_i, \qquad \text{for } j=p+2,\dots,d. \end{aligned} \tag{51}$$

Hence it remains to find $\det(d)$.

No matter the dimension, our basis elements are each one of seven types (in $\mathfrak{m}\cap\mathfrak{h}$, lower triangular in $\mathfrak{h}$, upper triangular in $\mathfrak{h}$, diagonal, in $\overline{\mathfrak{u}}$, in $\mathfrak{u}$, or in $\mathfrak{m}_1$). Depending on the type of $X_i$, we can calculate $d_{ij}$ explicitly independent of dimension via an inductive argument. From which it follows that the $d$ matrix has the following form:

$$d = \begin{pmatrix} I_{(n-1)(n-2)/2} & 0 & 0 & 0 & 0 \\ 0 & I_{n-1} & \frac{w^2}{2}I_{n-1} & 0 & wI_{n-1} \\ 0 & \frac{y^2}{2}I_{n-1} & \frac{1}{4}(2+wy)^2 I_{n-1} & 0 & \frac{1}{2}w(2+wy)I_{n-1} \\ 0 & 0 & 0 & B & 0 \\ 0 & 0 & 0 & 0 & I_{n-1} \end{pmatrix} \tag{52}$$

where $B$ is the $3\times 3$ matrix given by

$$B := \begin{pmatrix} 1+wy & -w & \frac{1}{2}(2+wy) \\ -y & 1 & -\frac{y^2}{2} \\ 0 & 0 & 1. \end{pmatrix}$$

Hence, one can determine explicitly thanks to a block matrix decomposition that $\det(d) = |1+wy|^{n-1}$. This completes the proof. □

## 4.5 Calculating the Casimir Operator

As in §2, given our basis, $X_1, X_2 \dots, X_d$, for the Lie algebra in §4.3, there is a corresponding dual basis $X_1^*, X_2^*, \dots X_d^*$ with respect to the Killing form. Then the Casimir operator can be written as the following second order differential operator:

$$\mathcal{C} := \sum_{i=1}^{p+2+\ell} X_i X_i^*. \tag{53}$$

Using the above argument, in $\mathbf{z}$ coordinates, we can express $X_i = \sum_{j=1}^{d} \eta_{ij}\partial_j$. Likewise, we can express $X_i^* = \sum_{j=1}^{d} \eta_{ij}^*\partial_j$.

Now define $\mu^*$ analogously to how we defined $\mu$, that is

$$\begin{aligned}\mathrm{Ad}((h_2\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_1 &= \sum_{i=1}^{d}\mu_{1i}^*X_i^*\\ \mathrm{Ad}((h_3\cdots h_p\bar{u}um_1\cdots m_\ell)^{-1})X_2 &= \sum_{i=1}^{d}\mu_{2i}^*X_i^*\\ &\cdots\\ X_d &= \sum_{i=1}^{d}\mu_{di}^*X_i^*.\end{aligned}$$

We then have that $\eta^*\cdot\mu^*\mathbf{X}^* = \mathbf{X}^*$, and hence $\eta^* = (\mu^*)^{-1}$. Thus, if we define $\underline{\partial} := (\partial_{x_1},\ldots,\partial_{n_d})^T$, then as a differential operator

$$\mathcal{C} = \left(\mu^{-1}\underline{\partial}\right)\cdot\left((\mu^*)^{-1}\underline{\partial}\right).$$

The Casimir operator in all $d$ variables is, of course, rather unwieldy. However, fortunately we shall only need the Casimir operator to act on left $H$-invariant, and right $M_1$-invariant functions. Hence, if we write $\mu^* = \nu^*c^*$, with $\nu^*$ and $c^*$ defined analogously to as above, then when acting on such functions, we have

$$\mathcal{C} = \left(c^{-1}\underline{\partial}\right)\cdot\left((c^*)^{-1}\underline{\partial}\right).$$

Using this decomposition, we can explicitly calculate the Casimir operator in $\mathbf{z}$ coordinates, when acting on left $H$-invariant and right $M_1$-invariant function.

**Theorem 54.** *[Structure Theorem for the Casimir Operator] Let $f : H\backslash G/M_1 \to \mathbb{C}$. Then in the $\mathbf{z}$-coordinate system, $f$ is only a function of $(y,w)$, that is, the $\overline{U}$ and $U$ variables, and the Casimir operator acting on $f$ has the following form:*

$$\mathcal{C}f(y,w) = -\frac{1}{2}\left(y^2\partial_y^2 + (n+1)y\partial_y + 2\partial_{yw} + \frac{(n-1)|w|\partial_w}{|1+yw|}\right)f(y,w). \tag{55}$$

*Proof.* By definition of $c^*$ we have

$$Y_j := \mathrm{Ad}((\bar{u}um_1\cdots m_\ell)^{-1})X_j = \sum_{i=1}^{d}c_{pi}^*X_i^* \text{ for } i = 1,\ldots,p$$

and

$$\begin{aligned}Y_{p+1} := \mathrm{Ad}((um_1\cdots m_\ell)^{-1})X_{p+1} &= \sum_{i=1}^{d}c_{(p+1)i}^*X_i^*\\ &\cdots\\ Y_d := X_d &= \sum_{i=1}^{d}c_{di}X_i^*.\end{aligned}$$

Thus, $(c^*)^{-1}\mathbf{Y} = \mathbf{X}^*$. Since our function depends only on $w$ and $y$, we are only interested in the $(p+1)^{th}$ and $(p+2)^{th}$ columns of $(c^*)^{-1}$. We can explicitly calculate $Y_i$ depending on which type of basis vector is $X_i$, then we can explicitly solve for $(c^*)^{-1}_{ij}$. Using an inductive argument we then have that $(c^*)^{-1}\underline{\partial}$ has the following explicit form: Let

$$v_M := (\sin(\varphi_1), \cos(\varphi_1)\sin(\varphi_2), \cos(\varphi_1)\cos(\varphi_2)\sin(\varphi_3), \ldots, \cos(\varphi_1)\cdots\cos(\varphi_{\ell-1})\sin(\varphi_\ell))^T \tag{56}$$

then we have (in fact, we only need the fourth and fifth row of the below)

$$(c^*)^{-1}\underline{\partial} = \frac{1}{2}\begin{pmatrix} 0_{(n-1)n/2} \\ -v_M\partial_w \\ \frac{1}{2}v_M(w^2\partial_w - 2(1+wy)\partial_y) \\ (y\partial_y - w\partial_w) \\ \cos(\varphi_1)\ldots\cos(\varphi_\ell)\partial_w \\ \frac{1}{2}\cos(\varphi_1)\ldots\cos(\varphi_\ell)(-w^2\partial_w + 2(1+wy)\partial_y) \\ 0_{n-2} \end{pmatrix}. \tag{57}$$

Let $(A_1,\ldots,A_d) := (c^*)^{-1}\cdot\underline{\partial}$, then our aim is to evaluate

$$\mathcal{C} = \sum_{j=1}^{d}\sum_{i=1}^{d}(c^{-1})_{ji}\partial_i A_j.$$

However, since we are considering the Casimir acting on right $M_1$ invariant functions, we may set the last $\ell$ coordinates equal to 0, that is

$$\mathcal{C}f(y,w) = \left(\sum_{j=1}^{d}\sum_{i=1}^{d}\left[(c^{-1})_{ji}\right]_{\varphi=0}\left[\partial_i A_j\right]_{\varphi=0}\right) f(y,w),$$

however $\left[(c^{-1})_{ji}\right]_{\varphi=0} = (d^{-1})_{ji}$ where $d$ is the matrix of coefficients defined in (51).

Since the $d$ matrix is explicitly described in (52), and the vector $(A_1,\ldots,A_d)$ is given explicitly in (57), then (55) can be found through direct computation. □

# 5 Counting

We turn now to the proof of Theorem 2. For simplicity, we assume the packing is bounded; similar methods apply for counting in regions. We also assume that the packing is the orbit of a single sphere $S_1$ under the action of a symmetry group $\Gamma$; in general, the counting problem reduces to a finite sum of such [KK23, Theorem 22]. Let the sphere $S_1$ be represented in the inversive coordinate system described in §4 by the vector $v_1 \in V$. We begin in the same way as we did in the $\mathrm{SL}_2(\mathbb{R})$ case, writing the count as

$$\begin{aligned} N_{\mathcal{P}}(T) &:= \#\{S \in \mathcal{P} \mid b(S) < T\} \\ &= \#\{v \in v_1\cdot\Gamma \mid b^*(v) < T\} \\ &= \sum_{\gamma\in\Gamma_1\backslash\Gamma}\chi_T(\gamma), \end{aligned}$$

where $b^*$ is the "bend" covector (as in (44)), $\Gamma_1 := \mathrm{Stab}_{v_1}(\Gamma)$, and for $g \in G$

$$\chi_T(g) := \begin{cases} 1 & \text{if } b^*(v_1 \cdot g) < T, \\ 0 & \text{else.} \end{cases} \tag{58}$$

We observe that $\chi_T$ is a function on $H\backslash G/UM_1$, where $H$ is the stabalizer of $v_1$ (so that $\Gamma_1 = \Gamma \cap H$).

Now we automorphize $\chi_T$, that is define

$$F_T(g) := \sum_{\gamma \in \Gamma_1 \backslash \Gamma} \chi_T(\gamma g).$$

Hence $F_T$ is left $\Gamma$-invariant and $N_{\mathcal{P}}(T) = F_T(e)$. Again, rather than trying to evaluate the discontinuous function $F_T$ at the origin, consider its inner product with a smooth approximation of the identity. However, we want to restrict as few directions of our smooth approximation as possible, to optimize the resulting error terms. To this end, fix an $\varepsilon > 0$, and let

$$\psi = \psi_\varepsilon \in L^2(\Gamma_1 \backslash G/M_1) \tag{59}$$

be given as follows. Let $\mathcal{F}$ be a fundamental domain described in $\mathbf{z}$-coordinates, of $\Gamma_1\backslash G/M_1$. Then on $\mathcal{F}$ we let $\psi$ be of the form

$$\psi(\mathbf{z}) = \psi_1(x_1) \cdots \psi_p(x_p) \psi_U(y) \psi_{\overline{U}}(w)$$

, with all components nonnegative and unit total mass,

$$\int_{\Gamma_1 \backslash G/M_1} \psi dg = 1,$$

and satisfying the following conditions.

For the $n+1$ variables in the components $H_+$, $H_A$ (in the notation of (46)) and $\overline{U}$, we restrict the coefficients to $\varepsilon$-balls around 0. The other variables are restricted only to compact regions of bounded size around 0. We can choose such a $\psi$ have $L^2$ mass bounded by:

$$\|\psi\|_{L^2(\Gamma_1 \backslash G/M_1)} \ll \varepsilon^{-(n+1)/2}.$$

Now, let $\mathcal{F}_0$ denote the fundamental domain for $\Gamma_1$ containing the origin, and define the function $w_T = w_{T,\varepsilon} : \Gamma_1 \backslash G \to [0,1]$ by

$$w_T(g) := \int_{\mathcal{F}_0} \chi_T(gh) \psi_\varepsilon(h) dh. \tag{60}$$

Our main counting theorem (Theorem 2) will follow from the following smooth effective count.

**Theorem 61.** *Let the Laplace eigenvalues of $\Gamma\backslash\mathbb{H}^{n+1}$ be as in (1). Then there exist constants $c^i_{\Gamma,\varepsilon}$ for $i = 0, 1, \ldots, k$, with $c^0_{\Gamma,\varepsilon} > 0$, such that*

$$\widetilde{N}_{\Gamma,\varepsilon}(T) := \sum_{\gamma \in \Gamma_1 \backslash \Gamma} w_T(\gamma) = c^0_{\Gamma,\varepsilon} T^\delta + c^1_{\Gamma,\varepsilon} T^{s_1} + \cdots + c^k_{\Gamma,\varepsilon} T^{s_k} + O\left( \frac{1}{\varepsilon^{(n+1)/2}} T^{n/2} \log(T) \right) \tag{62}$$

*where the implied constant depends only on $\Gamma$. Moreover $c^0_{\Gamma,\varepsilon} = c^0_\Gamma(1 + O(\varepsilon))$, and for $i \geq 1$ (if any exist), we have that $c^i_{\Gamma,\varepsilon} \ll \varepsilon^{-(n+1)/2}$.*

**Remark 63.** *Note that this smoothed counting theorem is optimal in that the error goes all the way to the tempered spectrum.*

## 5.1 Proof of Theorem 2

An explicit calculation shows that the inversive coordinates of a sphere which has been transformed by $g$ are

$$(y, -(1+wy)v_M^T, (1+wy)\cos(\varphi_1)\cdots\cos(\varphi_\ell), w(2+wy))$$

where $v_M$ is the vector defined in (56). Thus the 'bend radius' (i.e the distance of the center of the sphere to the origin multiplied by the bend) is $(1+wy)$. Since this distance is necessarily bounded from above (since the packing is bounded) and since $y$ is bounded from below (because it controls the bend, and the packing has a largest sphere), we know that $w$ is bounded from above. In particular we have

**Lemma 64.** *Let $\gamma \in \Gamma_1\backslash\Gamma$ then the $w$ coordinate of $\gamma$, $|w_\gamma|$ is bounded.*

Now, assuming Theorem 61 we present the proof of Theorem 2.

*Proof of Theorem 2.* Consider

$$w_T(\gamma) = \int_{\mathcal{F}_0} \chi_T(\gamma g)\psi_\varepsilon(g)dg.$$

We write $\gamma = \mathcal{J}(x_{1,\gamma}, \ldots, x_{k,\gamma}, y_\gamma, w_\gamma, \varphi_{1,\gamma}, \ldots, \varphi_{\ell,\gamma})$ and note that $\chi_T$ is left $H$-invariant and right $UM_1$ invariant. Hence

$$\chi_T(\gamma g) = \chi_T(\overline{u}(y_\gamma)u(w_\gamma)m_1(\varphi_\gamma)h(\mathbf{x})\overline{n}(y)).$$

now we can write $h$ as a product of one dimensional components of $H$ from (46) (here, it is convenient to change the order of the decomposition). That is, write

$$h(\mathbf{x}) = h_+(\mathbf{x}_+)h_A(\mathbf{x}_A)h_-(\mathbf{x}_-)h_M(\mathbf{x}_M).$$

Since $\psi$ restricts the $\mathbf{x}_-$ and $\mathbf{x}_A$ coordinates to balls of radius $\varepsilon$, we can apply adjoints to move those factors to the left of $\overline{u}(y_\gamma)$ and the other factors will be perturbed by an $\varepsilon$ error. Then we can use left $H$-invariance of $\chi_T$. We can do the same for the $\overline{u}(y)$ factor which gets absorbed in the $\overline{u}(y)$ factor. Thus, for $g$ in the support of $\psi_\varepsilon$

$$\begin{aligned}\chi_T(\gamma g) &= \chi_T(\overline{u}(y_\gamma + O(\varepsilon))u(w_\gamma + O(\varepsilon))m_1(\varphi_\gamma + O(\varepsilon))h_+(\mathbf{x}_+ + O(\varepsilon))h_M(\mathbf{x}_M + O(\varepsilon)) \\ &= \chi_T(\overline{u}(y_\gamma + O(\varepsilon))),\end{aligned}$$

where the second line follows from right-$M$ invariance, and right-$H_+$ invariance.

Thus,

$$\chi_T(\gamma g) = \begin{cases} 1 & \text{if } y_\gamma < \frac{T}{1+c\varepsilon} \\ 0 & \text{if } y_\gamma > \frac{T}{1-c\varepsilon} \end{cases}$$

for some absolute constant $c$.

Hence, since $\widetilde{N}_{\Gamma,\varepsilon}(T) = \sum_{\Gamma_1\backslash\Gamma} w_T(\gamma)$, then our count satisfies:

$$\widetilde{N}_{\Gamma,\varepsilon}(T(1-c\varepsilon)) \le N_{\mathcal{P}}(T) \le \widetilde{N}_{\Gamma,\varepsilon}(T(1+c\varepsilon)) \tag{65}$$

Now, assuming there are no other eigenvalues, we apply Theorem 61 to find:

$$\widetilde{N}_{\Gamma,\varepsilon}(T(1\pm\varepsilon)) = c^0(1+O(\varepsilon))T^\delta + O(\frac{1}{\varepsilon^{(n+1)/2}}T^{n/2}\log(T)).$$

Then choosing $\varepsilon = T^{\frac{2}{n+3}(n/2-\delta)}\log(T)^{2/(n+3)}$ optimizes this inequality. Thus

$$\widetilde{N}_{\Gamma,\varepsilon}(T(1\pm\varepsilon)) = c^0 T^\delta + O(T^{\frac{2}{n+3}(n/2-\delta)+\delta}\log(T)^{\frac{2}{n+3}}).$$

The general case follows similarly. □

## 5.2 Inserting the Casimir Operator

Once again the smooth count is the inner product of $F_T$ with an $L^2$ function. Consider the inner product of $F_T$ with a general $\Psi \in L^2(\Gamma\backslash G/M_1)$. That is, given a function $\psi$ on $\Gamma_1\backslash G/M_1$ we automorphize it

$$\Psi(z) := \sum_{\gamma\in\Gamma_1\backslash\Gamma} \psi(\gamma g).$$

Let

$$K_T(s) := \frac{T^s b^{n-s} - T^{n-s}b^s}{b^{n-s}-b^s}, \qquad L_T(s) := \frac{T^{n-s}-T^s}{b^{n-s}-b^s},$$

while for $s = n/2 + it$ we have

$$K_T(s) := T^{n/2}\frac{\sin(t\log T/\log b)}{\sin(t\log b)}, \qquad L_T(s) := \left(\frac{T}{b}\right)^{n/2}\frac{\sin(t\log T)}{\sin(t\log b)}. \tag{66}$$

Again, by choosing an appropriate choice of $b$ one can ensure that

$$K_T(s), L_T(s) \ll \begin{cases} T^s & \text{if } s\in(n/2,n], \\ T^{n/2}\log T & \text{if } s = n/2+it. \end{cases} \tag{67}$$

In the $\mathrm{SL}_2(\mathbb{R})$ case, we decomposed the real direction, since $\Gamma\backslash\mathbb{H}$ was not compact in the real direction. Analogously in the current setting, we have the group decomposition $(\Gamma_1\backslash H)\overline{U}UM_1$. Since $(\Gamma_1\backslash H)$ has finite $H$-Haar measure, and $M_1$ is compact, and we have imposed a cut-off in the $\overline{U}$-direction, we are again faced with a one dimensional non-compact direction, the $U$-direction. To that end, since $\infty$ lies outside the limit set, from Lemma 64 it follows that there exists an $X$ large enough, such that

$$N_{\mathcal{P}} = F_{T,X}(e)$$

where

$$F_{T,X}(g) := \sum_{\Gamma_1\backslash\Gamma} \chi_T(\gamma g)\widetilde{\chi}_X(\gamma g), \qquad \text{and}$$
$$\widetilde{\chi}_X(g) := \begin{cases} 1 & \text{if } |bz^*(v_1\cdot g)|/b(c_1\cdot g) < X, \\ 0 & \text{otherwise.} \end{cases}$$

where $bz^*$ denotes the bend center, as in (45) and $b$ denotes the bend. Thus, note that $\widetilde{\chi}_X$ is also left $\Gamma_1$ invariant, and right $M_1$ invariant (since it depends only on the distance to the origin of the center, not the polar coordinate angles).

Moreover, a calculation shows that, in the $\mathbf{z}$-coordinates, $\widetilde{\chi}_X$ can be written as:

$$\widetilde{\chi}_X(g) := \begin{cases} 1 & \text{if } -X - 1/y < w < X - 1/y, \\ 0 & \text{otherwise.} \end{cases} \tag{68}$$

**The Difference Operator:** Again, we will prove an identity in terms of $K_T$ and $L_T$ for $F_{T,X}$. To that end, consider the difference operator

$$G_{T,X} := F_{T,X} - K_T(\mathcal{C})F_{1,X} - L_T(\mathcal{C})F_{b,X}.$$

By self-adjointness of $\mathcal{C}$, for any $\Psi \in L^2(\Gamma\backslash G/M_1)$ we have

$$\begin{aligned} G_{T,X}(\Psi) &:= \langle G_{T,X}, \Psi\rangle_{\Gamma\backslash G} \\ &= \int_{\Gamma\backslash G} [F_{T,X}(g)\Psi(g) - F_{1,X}(g)(K_T(\mathcal{C})\Psi)(g) - F_{b,X}(g)(L_T(\mathcal{C})\Psi)(g)]\, dg, \end{aligned}$$

which we can unfold to

$$\begin{aligned} &G_{T,X}(\Psi) = \\ &\int_{\Gamma_1\backslash G} \widetilde{\chi}_X(g)\,(\chi_T(g)\Psi(g) - \chi_1(g)(K_T(\mathcal{C})\Psi)(g) - \chi_b(g)(L_T(\mathcal{C})\Psi)(g))\, dg. \end{aligned} \tag{69}$$

It is more convenient to work using $\mathbf{z}$-coordinates describing the group $G$. Now fix a fundamental domain for $\Gamma_1\backslash G/M_1$, this fundamental domain can be written in coordinates as $\mathcal{F} := P_{\Gamma_1} \times [0,\infty) \times \mathbb{R} \times [-\pi,\pi]^{n-1}$, where $P_{\Gamma_1}$ is a (finite $H$-volume) fundamental domain for the action of $\Gamma_1$ on $H$. Given a function $f : G \to \mathbb{C}$ we abuse notation and write $f(\mathbf{z}) = f(g_{\mathbf{z}})$, thus if we let $\rho$ denote density of the Haar measure, we can write (69) as:

$$\begin{aligned} &G_{T,X}(\Psi) = \\ &\int_{\mathcal{F}} \widetilde{\chi}_X(y,w)\,(\chi_T(y)\Psi(\mathbf{z}) - \chi_1(y)(K_T(\mathcal{C})\Psi)(\mathbf{z}) - \chi_b(y)(L_T(\mathcal{C})\Psi)(\mathbf{z}))\,\rho(\mathbf{z})d\mathbf{z}. \end{aligned}$$

Note that by definition $\chi_T(\mathbf{z})$ is a function of $y$ and $\widetilde{\chi}_X(\mathbf{z})$ is a function of $w$ (and $y$), for clarity we ignore the dependence on the other variables.

Let $\mathcal{F}_X := \{z \in \mathcal{F} \,:\, w \in \mathcal{I}_X\}$ where $\mathcal{I}_X := [w_{-X}(y), w_X(y)]$; here we have introduced the notation

$$w_r(y) := r - 1/y,$$

motivated by (68). Now write

$$G_{T,X}(\Psi) = \int_{\mathcal{F}_X} (\chi_T(y)\Psi(\mathbf{z}) - \chi_1(y)(K_T(\mathcal{C})\Psi)(\mathbf{z}) - \chi_b(y)(L_T(\mathcal{C})\Psi)(\mathbf{z}))\,\rho(\mathbf{z})d\mathbf{z}.$$

By showing that $G_{T,X}(\Psi) = 0$ for any choice of $\Psi$ we will prove the following.

**Proposition 70.** *For $\Gamma$ and $F_{T,X}$ as above we have that*

$$F_{T,X} = K_T(\mathcal{C})F_{1,X} + L_T(\mathcal{C})F_{b,X} \tag{71}$$

*where $K_T$ and $L_T$ are the differential operators defined in* (66).

Once again our goal is to work on the fundamental domain of a group (rather than working with discontinuous cut-offs). Thus we will perform the same smoothing as we did in the $\mathrm{SL}_2(\mathbb{R})$ case.

**Smoothing $\chi_T$:** Let $\sigma > 0$ and let

$$\chi_{1,\sigma}(\mathbf{z}) := \begin{cases} 1 & \text{if } y < 1, \\ 0 & \text{if } y > (1+\sigma), \end{cases}$$

and let $\chi_{1,\sigma}$ interpolate smoothly for all values in between. Now let $\chi_{T,\sigma}(\mathbf{z}) := \chi_{1,\sigma}(Ty)$.

Let

$$G^{\sigma}_{T,X}(\Psi) := \int_{\mathcal{F}_X} (\chi_{T,\sigma}(y)\Psi(\mathbf{z}) - \chi_{1,\sigma}(y)(K_T(\mathcal{C})\Psi)(\mathbf{z}) - \chi_{b,\sigma}(y)(L_T(\mathcal{C})\Psi)(\mathbf{z}))\,\rho(\mathbf{z})d\mathbf{z}.$$

Now by construction

$$\int_{\mathcal{F}_X} |\chi_{t,\sigma}(y) - \chi_t(y)|^2 \rho(\mathbf{z})d\mathbf{z} \ll_{t,X} \sigma$$

Thus by Cauchy-Schwarz we have that $\lim_{\sigma\to 0} G^{\sigma}_{T,X}(\Psi) = G_{T,X}(\Psi)$. Now our goal is to show that for any fixed $\varepsilon > 0$ we have $G^{\sigma}_{T,X}(\Psi) < \varepsilon$.

**Periodizing and Smoothing $\Psi$:** Let

$$\mathcal{I}_{X,\eta} := [w_{-X}(y) + \eta, w_X(y) - \eta].$$

Let $\widetilde{\Psi} : G/M_1 \to \mathbb{R}$ denote a function which agrees with $\Psi$ on

$$\{z \in \mathcal{F} \,:\, w \in \mathcal{I}_{X,\eta}\}$$

for some $\eta > 0$ to be chosen later. When $w \notin \mathcal{I}_{X,\eta/2}$ we impose the condition $\widetilde{\Psi}(w_{-X}(y)) = \widetilde{\Psi}(w_X(y))$ for any value of the other variables, and interpolate smoothly in between.

Using the same Cauchy-Schwarz argument as we employed in the $\mathrm{SL}_2(\mathbb{R})$ case, we can choose $\eta$ such that the $L^2(\mathcal{F}_X)$ cost of moving from $\Psi$ to $\widetilde{\Psi}$ is less than $\varepsilon$. Thus, Proposition 70 follows if we can prove that

$$G_{T,X,\sigma}(\widetilde{\Psi}) = 0. \tag{72}$$

**Working on $\mathcal{F}_X$:** Let

$$g^{\sigma}_{T,X}(\mathbf{z}) := \widetilde{\chi}_X(\mathbf{z})(\chi_{T,\sigma}(\mathbf{z}) - K_T(\mathcal{C})\chi_{1,\sigma}(\mathbf{z}) - L_T(\mathcal{C})\chi_{b,\sigma}(\mathbf{z})).$$

Furthermore, let $L^2(\mathcal{F}_X)$ denote the space of square integrable functions on $\mathcal{F}_X$, which are $\Gamma_1$-invariant in the $H$ variables, and invariant under translation by $2X$ in the $w$ variable. Note that $\widetilde{\Psi} \in L^2(\mathcal{F}_X)$. Then (72) follows from the following lemma.

**Lemma 73.** *For any $\psi \in L^2(\mathcal{F}_X)$ which is independent of the $\varphi \in [-\pi,\pi)^{n-1}$ variables, and any $\lambda \geq 0$, we have*

$$\langle g^{\sigma}_{T,X}, \psi\rangle_{\mathcal{F}_X} \ll_{\lambda,T,\sigma,X} \|(\mathcal{C}-\lambda)\psi\|_{\mathcal{F}_X} \tag{74}$$

*where $\mathcal{C}$ denotes the Casimir operator on $L^2(\mathcal{F}_X)$ (i.e the Casimir operator in $\mathbf{z}$-coordinates).*

*Proof.* To prove (74), let $\psi$ be an arbitrary function in $L^2(\mathcal{F}_X)$, which is $\varphi$ invariant. Now consider

$$\int_{\mathcal{F}_X} \chi_{T,\sigma}(y)\psi(\mathbf{z})\rho(\mathbf{z})d\mathbf{z} = \int_0^{2T} \chi_{T,\sigma}(y) \int_{w_{-X}(y)}^{w_X(y)} \int_{[-\pi,\pi)^{n-1}} \int_{P_{\Gamma_1}} \psi(\mathbf{z})\rho(\mathbf{z})d\mathbf{x}d\varphi dw dy.$$

Since, as shown in Theorem 47, the density of the Haar measure decomposes into a product of densities depending on $\mathbf{x}$, one depending on $\varphi$ and one depending on $w$ and $z$, we can use this fact to show

$$\int_{P_{\Gamma_1}} \mathcal{C}\psi(\mathbf{x},y,w)\rho_H(\mathbf{x})d\mathbf{x} = \mathcal{C}\int_{P_{\Gamma_1}} \psi(\mathbf{x},y,w)\rho_H(\mathbf{x})d\mathbf{x}.$$

This follows from the fact that $\psi$ is $\Gamma_1$-invariant in the $\mathbf{x}$ variable, and $P_{\Gamma_1}$ is a fundamental domain for $\Gamma_1$ acting on $H$.

Recall that the Haar measure, in $\mathbf{z}$ coordinates is given by $|1+wy|^{n-1}\rho_H\rho_{M_1}$. Let $\widetilde{\psi}(y,w) := \int_{P_{\Gamma_1}} \psi(\mathbf{x},y,w)\rho_H(\mathbf{x})d\mathbf{x}$ and let

$$h(y,w) := 2\int_{P_{\Gamma_1}} \int_{[-\pi,\pi)^{n-1}} \widetilde{\rho}(\mathbf{z})(\mathcal{C}-\lambda)\psi(\mathbf{z})d\mathbf{x}d\varphi.$$

Then, using the Structure Theorem for the Casimir operator (Theorem 54), we have that

$$\begin{aligned} I(y) &:= \int_{w_{-X}(y)}^{w_X(y)} h(w,y)dw \\ &= -\int_{w_{-X}(y)}^{w_X(y)} |1+wy|^{n-1}\left(y^2\partial_y^2 + (n+1)y\partial_y + 2\partial_{yw} + \frac{(n-1)w\partial_w}{1+yw} + \lambda\right)\widetilde{\psi}(y,w)dw \end{aligned} \tag{75}$$

Now make the change of variables $w \mapsto w' = w + 1/y$ and $y' = y$. This leads to the relations

$$\partial_w = \partial_{w'}, \qquad \text{and} \qquad \partial_y = \partial_{y'} - \frac{1}{y^2}\partial_{w'},$$

and results in

$$I(y) = -\int_{-X}^{X} y^{n-1}\,|w|^{n-1}\left(y^2\partial_{yy} + (n+1)y\partial_y - \frac{1}{y^2}\partial_{ww} - \frac{(n-1)}{wy^2}\partial_w + \lambda\right)\widetilde{\psi}(y,w)dw.$$

Now, we can use integration by parts to eliminate two of the terms

$$I(y) = -y^{n-1}\int_{-X}^{X} |w|^{n-1}\left(y^2\partial_{yy} + (n+1)y\partial_y + \lambda\right)\widetilde{\psi}(y,w)dw.$$

We now let $f(y) := \int_{-X}^{X} |w|^{n-1}\,\widetilde{\psi}(y,w)dw$; then $f(y)$ satisfies the differential equation

$$\left(y^2\partial_{yy} + (n+1)y\partial_y + \lambda\right) f(y) = h(y) \tag{76}$$

with $h(y) = -\int_{w_{-X}(y)}^{w_X(y)} h(w,y)dw$.

The homogeneous solution to this differential equation is then

$$C_1 y^{s-n} + C_2 y^{-s},$$

where we write $\lambda = s(n-s)$. From here we can use the usual variation of parameters argument to conclude

$$\int_0^T |y|^{n-1}\int_{-X}^{X}\widetilde{\psi}dwdy = C_1T^s + C_2T^{n-s} + O_{T,\lambda,\Gamma}(\|(\mathcal{C}-\lambda)\widetilde{\psi}\|).$$

□

Now, with Lemma 73 we use the same argument as in [Kon09, Proof of Theorem 3.2] to conclude that $g^{\sigma}_{T,X}$ is almost everywhere 0. Thus, working our way back up, with the same argument as in the $\mathrm{SL}_2(\mathbb{R})$ setting, we conclude the proof of Proposition 70.

## 5.3 Proof of Theorem 61

As for the $SL_2(\mathbb{R})$ case, we now return to the smooth count with $U$-cutoff:

$$\widetilde{N}_\Gamma(T) = \langle F_{T,X}, \Psi\rangle_\Gamma.$$

Now apply the abstract Parseval's identity (13)

$$\begin{aligned}\langle F_{T,X}, \Psi\rangle_\Gamma &= \langle \widehat{F_{T,X}}, \widehat{\Psi}\rangle_{\mathrm{Spec}(\Gamma)} \\ &= \widehat{F_{T,X}}(\lambda_0)\widehat{\Psi}(\lambda_0) + \int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \widehat{F_{T,X}}(\lambda)\widehat{\Psi}(\lambda)\mathrm{d}\nu(\lambda).\end{aligned} \tag{77}$$

As with the $\mathrm{SL}_2(\mathbb{R})$ case, we can apply spectral methods to extract the $T$ and $\varepsilon$-dependence. Applying the abstract spectral theorem gives

$$\widehat{F_{T,X}}(\lambda_0)\widehat{\Psi}(\lambda_0) = \langle \mathrm{Proj}_{\mathcal{H}_0}(F_{T,X}), \mathrm{Proj}_{\mathcal{H}_0}(\Psi)\rangle.$$

Then by linearity and our main identity Proposition 70 we conclude

$$\begin{aligned}\widehat{F_{T,X}}(\lambda_0)\widehat{\Psi_\varepsilon}(\lambda_0) &= K_T(\lambda_0)\langle \mathrm{Proj}_{\mathcal{H}_0}(F_{1,X}), \mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)\rangle \\ &\quad + L_T(\lambda_0)\langle \mathrm{Proj}_{\mathcal{H}_0}(F_{b,X}), \mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)\rangle \\ &= T^\delta \langle H, \mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)\rangle + O(T^{n/2}),\end{aligned}$$

where $H := c_1 \mathrm{Proj}_{\mathcal{H}_0}(F_{1,X}) + c_b \mathrm{Proj}_{\mathcal{H}_0}(F_{b,X})$ for some constants $c_1, c_b$.

The problem remains to determine the $\varepsilon$ dependence of the projection operator $\mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)$. In general, this projection can be realized in a number of ways, either as a Burger-Roblin-type measure of $\Psi_\varepsilon$ (see [MO15, p. 861]), or using representation-theoretic decompositions as in [BKS10, Vin12]. We will give a soft argument that avoids either.

We know from (65) that

$$N_{\mathcal{P}}\left(\frac{T}{1+c\varepsilon}\right) \le \widetilde{N}_{\Gamma,\varepsilon}(T) \le N_{\mathcal{P}}\left(\frac{T}{1-c\varepsilon}\right),$$

for any value of $\varepsilon$ and any value of $T$. However we also know *a priori* (e.g., using [Kim15]) that

$$N_{\mathcal{P}}\left(\frac{T}{1\pm c\varepsilon}\right) = c_{\mathcal{P}}\left(\frac{T}{1\pm c\varepsilon}\right)^\delta (1+o(1)),$$

as $T\to\infty$. Dividing by $T^\delta$ then gives

$$c_{\mathcal{P}}\left(\frac{1}{1+c\varepsilon}\right)^\delta - o(1) \le \langle H, \mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)\rangle \le c_{\mathcal{P}}\left(\frac{1}{1-c\varepsilon}\right)^\delta + o(1).$$

Now send $T\to\infty$ and Taylor expand $\frac{1}{(1\pm c\varepsilon)^\delta}$ in $\varepsilon$, giving:

$$\langle H, \mathrm{Proj}_{\mathcal{H}_0}(\Psi_\varepsilon)\rangle = C + O(\varepsilon).$$

Hence

$$\widehat{F_{T,X}}(\lambda_0)\widehat{\Psi_\varepsilon}(\lambda_0) = T^\delta c(1+O(\varepsilon)) + O(T^{n/2}) \tag{78}$$

for some constant $c$ independent of $\varepsilon$. (Note that this positivity argument does not apply to the other eigenvalues. Hence with sharp cutoffs, as in Theorem 2, we do not extract lower order terms.)

Turning now to the remainder, after extracting the main term corresponding to $\lambda_0$ we are left with

$$\begin{aligned}\mathrm{Err} &:= \int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \widehat{F_{T,X}}(\lambda)\widehat{\Psi}(\lambda)d\nu \\ &= \int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \left(K_T(\lambda)\widehat{F_{1,X}}(\lambda) + L_T(\lambda)\widehat{F_{b,X}}(\lambda)\right)\widehat{\Psi}(\lambda)d\nu.\end{aligned}$$

Assume for simplicity that there are no other discrete eigenvalues above the base. Now apply the abstract spectral theorem and the bounds from (67) to conclude that

$$\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} K_T(\lambda)\widehat{F_{1,X}}(\lambda)\widehat{\Psi}(\lambda)d\nu \ll T^{n/2}\log T \int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \widehat{F_{1,X}}(\lambda)\widehat{\Psi}(\lambda)d\nu.$$

Now apply Cauchy-Schwarz and Parseval to get

$$\begin{aligned}
&\ll T^{n/2}\log T \left(\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \widehat{F_{1,X}}(\lambda)^2 d\nu\right)^{1/2} \left(\int_{\mathrm{Spec}(\Gamma)\setminus\{\lambda_0\}} \widehat{\Psi}(\lambda)^2 d\nu\right)^{1/2} \\
&\leq T^{n/2}\log T \left(\int_{\mathrm{Spec}(\Gamma)} \widehat{F_{1,X}}(\lambda)^2 d\nu\right)^{1/2} \left(\int_{\mathrm{Spec}(\Gamma)} \widehat{\Psi}(\lambda)^2 d\nu\right)^{1/2} \\
&= T^{n/2}\log T \; \|F_{1,X}\|_\Gamma \, \|\Psi\|_\Gamma.
\end{aligned}$$

Finally, note that since $\psi_\varepsilon$ is normalized to have unit $L^1$-mass, we have that $\|\Psi\|_\Gamma \ll \varepsilon^{-(n+1)/2}$. In the case of other eigenvalues, we replace the bound $T^{n/2}\log T$ above with $T^{s_1}$. This completes the proof of Theorem 61. □

**Remark 79.** *If we remove our assumption that $L^2(\Gamma\backslash G)$ does not weakly contain any nonspherical complementary series representations, then, after removing contributions from Laplace eigenvalues in (77), the remainder would not necessarily be tempered. So instead of getting an error of order $T^{n/2}\log T$, we would only be able to bound what remains by $O(T^{n-1})$, corresponding to the spectral parameter of any potential nonspherical complimentary series. It is likely possible to improve on this bound, see Remark 7.*

## Acknowledgments

The authors are grateful to Sam Edwards and Stephen D. Miller for enlightening conversations and suggestions, and to Iván Rasskin for allowing us to use his image in Figure 1 generated from his `polytopack` Mathematica package (in preparation). Thanks also to Hee Oh for comments on an earlier draft, and to the referees for their detailed comments and corrections.

## References

[BKS10] J. Bourgain, A. Kontorovich, and P. Sarnak. Sector estimates for hyperbolic isometries. *Geometric and Functional Analysis*, 20(5):1175–1200, Nov 2010. 49

[Boy73] D. Boyd. The residual set dimension of the Apollonian packing. *Mathematika*, 20:170–174, 1973. 16

[EO21] S. Edwards and H. Oh. Spectral gap and exponential mixing on geometrically finite hyperbolic manifolds. *Duke Math. J.*, 170(15):3417–3458, 2021. 16, 17

[Kim15] I. Kim. Counting, mixing and equidistribution of horospheres in geometrically finite rank one locally symmetric manifolds. *J. Reine Angew. Math.*, 704:85–133, 2015. 16, 49

[KK23] M. Kapovich and A. Kontorovich. On superintegral kleinian sphere packings, bugs, and arithmetic groups. *J. Reine Angew. Math.*, 798:105–142, 2023. 15, 20, 31, 41

[KKL23] D. Kelmer, A. Kontorovich, and C. Lutsko. Sarnak's spectral gap question. *Journal d'Analyse Mathematique*, 151:171–179, 2023. 17

[KN19] A. Kontorovich and K. Nakamura. Geometry and arithmetic of crystallographic sphere packings. *Proc. Natl. Acad. Sci. USA*, 116(2):436–441, 2019. 20

[Kna01] A. Knapp. *Representation theory of semisimple groups*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001. An overview based on examples, Reprint of the 1986 original. 21

[KO11] A. Kontorovich and H. Oh. Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds. *J. Amer. Math. Soc.*, 24(3):603–648, 2011. With an appendix by Oh and Nimish Shah. 16

[KO12] A. Kontorovich and H. Oh. Almost prime Pythagorean triples in thin orbits. *Journal fur die Reine und Angewandte Mathematik*, 667:89–131, 2012. 19, 22

[Kon09] A. Kontorovich. The hyperbolic lattice point count in infinite volume with applications to sieves. *Duke Math. J.*, 149(1):1–36, 2009. 15, 16, 17, 18, 19, 20, 22, 24, 28, 29, 30, 48

[Kon13] A. Kontorovich. From Apollonius to Zaremba: local-global phenomena in thin orbits. *Bull. Amer. Math. Soc. (N.S.)*, 50(2):187–228, 2013. 15

[LO13] M. Lee and H. Oh. Effective circle count for Apollonian packings and closed horospheres. *Geom. Funct. Anal.*, 23(2):580–621, 2013. 16, 17

[LP82] P. Lax and R. Phillips. The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces. *J. Functional Analysis*, 46(3):280–350, 1982. 15, 16, 18, 22

[MO15] A. Mohammadi and H. Oh. Matrix coefficients, counting and primes for orbits of geometrically finite groups. *Journal of the EMS*, 17:837–897, 2015. 16, 49

[Pan17] W. Pan. Effective equidistribution of circles in the limit sets of Kleinian groups. *J. Mod. Dyn.*, 11:189–217, 2017. 16

[Pat76] S. Patterson. The limit set of a Fuchsian group. *Acta Mathematica*, 136:241–273, 1976. 16, 22

[Rud73] W. Rudin. *Functional analysis*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973. 21

[Sar07] P. Sarnak. Letter to J. Lagarias, 2007. http://web.math.princeton.edu/sarnak/AppolonianPackings.pdf. 17

[Sul84] D. Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153:259–277, 1984. 16

[Vin12] I. Vinogradov. *Effective bisector estimate with application to Apollonian circle packings*. ProQuest LLC, Ann Arbor, MI, 2012. Thesis (Ph.D.)–Princeton University. 16, 49

## AUTHORS

Alex Kontorovich
Department of Mathematics,
Rutgers University,
New Brunswick, NJ, USA
alex [dot] kontorovich [at] rutgers [dot] edu
https://math.rutgers.edu/~alexk/

Christopher Lutsko
Department of Mathematics,
Rutgers University,
New Brunswick, NJ, USA
chris [dot] lutsko [at] rutgers [dot] edu
https://chrislutsko.com/