STRONG SPECTRAL GAPS FOR COMPACT QUOTIENTS OF PRODUCTS OF $\mathrm{PSL}(2,\mathbb{R})$

DUBI KELMER AND PETER SARNAK

Abstract. The existence of a strong spectral gap for quotients $\Gamma\backslash G$ of noncompact connected semisimple Lie groups is crucial in many applications. For congruence lattices there are uniform and very good bounds for the spectral gap coming from the known bounds towards the Ramanujan-Selberg Conjectures. If G has no compact factors then for general lattices a strong spectral gap can still be established, however, there is no uniformity and no effective bounds are known. This note is concerned with the strong spectral gap for an irreducible co-compact lattice Γ in $G = \mathrm{PSL}(2,\mathbb{R})^d$ for $d \geq 2$ which is the simplest and most basic case where the congruence subgroup property is not known. The method used here gives effective bounds for the spectral gap in this setting.

INTRODUCTION

This note is concerned with the strong spectral gap property for an irreducible co-compact lattice Γ in $G = \mathrm{PSL}(2,\mathbb{R})^d$, $d \geq 2$. Before stating our main result we review in some detail what is known about such spectral gaps more generally. Let G be a noncompact connected semisimple Lie group with finite center and let Γ be a lattice in G. For π an irreducible unitary representation of G on a Hilbert space H, we let $p(\pi)$ be the infimum of all p such that there is a dense set of vectors $v \in H$ with $\langle \pi(g)v, v\rangle$ in $L^p(G)$. Thus if π is finite dimensional $p(\pi) = \infty$, while π is tempered if and only if $p(\pi) = 2$. In general $p(\pi)$ can be computed from the Langlands parameters of π and for many purposes it is a suitable measure of the non-temperedness of π (if $p(\pi) > 2$). The regular representation, $f(x) \mapsto f(xg)$, of G on $L^2(\Gamma\backslash G)$ is unitary and if $\Gamma\backslash G$ is compact it decomposes into a discrete direct sum of irreducibles while if $\Gamma\backslash G$ is non-compact the decomposition involves also continuous integrals via Eisenstein series. In any case, let

Date: July 1, 2010.

The first author was partially supported by the NSF grant DMS-0635607, and the second author by the NSF grant DMS-0758299.

E denote the exceptional exponent set defined by

$$E(\Gamma\backslash G) = \left\{ p(\pi) : \begin{array}{c} p(\pi) > 2, \text{ and } \pi \text{ is an infinite} \\ \text{dimensional irreducible representation} \\ \text{of } G \text{ occurring weakly in } L^2(\Gamma\backslash G) \end{array} \right\} \tag{0.1}$$

If $E(\Gamma\backslash G)$ is empty set $p(\Gamma\backslash G) = 2$ and otherwise let

$$p(\Gamma\backslash G) = \sup E(\Gamma\backslash G) \tag{0.2}$$

We say that $\Gamma\backslash G$ has a strong spectral gap if $p(\Gamma\backslash G) < \infty$. The existence of such a gap is critical in many applications. In the case that Γ is a congruence group (this is the automorphic case discussed below) the set E and the precise value of p are closely connected to the generalized Ramanujan conjectures [8, 40]. In ergodic theoretic applications $p(\Gamma\backslash G)$ controls the precise mixing rate of the action of noncompact subgroups of G on $\Gamma\backslash G$ [36, 19, 42]. In questions of local rigidity of related actions the spectral gap controls the "small divisors" in the linearized co-cycle equations [12] and it plays an important role in the study of the cohomology of Γ [3, 4].

The congruence case is defined as follows. Let H be a semisimple linear algebraic group defined over a number field F and let S_∞ denote the set of archimedean places of F. The group $G = \prod_{\nu\in S_\infty} H(F_\nu)$, where F_ν is the completion of F at ν, and Γ is a congruence subgroup of $H(F)$ embedded into G diagonally. After [7] and [9], which establish bounds towards Ramanujan Conjectures in general, one knows that $p(\Gamma\backslash G)$ is finite in these cases. In fact the methods used there yield explicit, and in many cases quite sharp, bounds for $p(\Gamma\backslash G)$ which depend only on H and not on Γ. The latter is crucial in many number theoretical as well as group theoretic applications [33, 40]. Arthur's conjectures [1, 8] for the discrete spectrum for such spaces $\Gamma\backslash G$ imply strong restrictions on the non-tempered π's that can occur. Specifically they must correspond to local Arthur parameters which gives a "purity" property [3, Chapter 6] and which in turn restricts the set $E(\Gamma\backslash G)$. In particular the set $E(\Gamma\backslash G)$ should be finite, though the set of non-tempered π's occurring in $L^2(\Gamma\backslash G)$ can certainly be infinite.

Two basic congruence examples are (i) $G = \mathrm{SL}(2,\mathbb{R})$ and Γ a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$ in which case Selbergs eigenvalue Conjecture [44] is equivalent to $E(\Gamma\backslash G) = \emptyset$, while it is known that $E(\Gamma\backslash G)$ is finite and is contained in $(2, \frac{64}{25}]$ [23]. (ii) $G = \mathrm{SL}(3,\mathbb{R})$ and Γ a congruence subgroup of $\mathrm{SL}(3,\mathbb{Z})$ in which case Langlands Conjectures for automorphic cuspidal representations on GL_n [26] imply that $E(\Gamma\backslash G) = \{4\}$. The exceptional exponent comes from the unitary Eisenstein series for the maximal parabolic subgroup. From [23] it follows that $E(\Gamma\backslash G) \subset \{4\} \cup (2, \frac{28}{9}]$ but here E is not known to be finite.

Returning to the general lattice Γ, we may, without any serious loss of generality take G to be the direct product $G_1 \times G_2 \times \cdots \times G_n$ of simple Lie groups with trivial center and assume that Γ is irreducible. By the latter we mean that if $G = G_1 \times G_2 \cdots \times G_r \times G_c$ with G_j noncompact for $j = 1, \ldots, r$ and G_c compact, that the projection of Γ onto the compact factor is dense and if $r > 1$ so are the projections of Γ on each $G_j,\ j = 1, \ldots, r$. This implies that the only G_j invariant vector in $L^2(\Gamma\backslash G)$ is the constant function.

If G has no compact factors then $L^2(\Gamma\backslash G)$ has a strong spectral gap. To see this consider separately the case that $r = 1$ and $r > 1$. If $r = 1$ and the rank of G is 1, then the spectral gap follows directly from the discreteness of the spectrum of the Laplacian below the (possible) continuous spectrum on $\Gamma\backslash G/K$ where K is a maximal compact subgroup of G (in the cases $G = \mathrm{SO}(n, 1)$ and $\mathrm{SU}(n, 1)$, $p(\Gamma\backslash G)$ can be arbitrarily large as one varies over Γ, this is shown for $\mathrm{SO}(2, 1)$ in [44] by starting with a Γ with $H_1(\Gamma)$ infinite and this can be done in the same way for these G's). If the rank of G is at least 2 then G has property T and $p(\Gamma\backslash G)$ is less than or equal to $p(G)$ which is finite [11]. The optimal exponent $p(G)$ associated with such a G has been determined in many cases including classical groups [18, 28] and some exceptional groups [30], while explicit and strong upper bounds for $p(G)$ are given for split exceptional groups in [29] and in complete generality in [39]. If $r \geq 2$, we need to use more machinery to deduce the spectral gap. Firstly by Margulis [35, Capter IX], Γ is arithmetic and hence is commensurable with a congruence lattice of the type discussed in the previous paragraph, for which we have a strong spectral gap. This coupled with the Lemma of Furman-Shalom, Kleinbock-Margulis (see [25, page 462]) allows one to pass from the congruence group to Γ and to conclude that $p(\Gamma\backslash G) < \infty$. Note that any π occurring in $L^2(\Gamma\backslash G)$ is of the form $\pi \cong \pi_1 \otimes \pi_2 \cdots \otimes \pi_r$, with π_j an irreducible representation of G_j and that $p(\pi) = \max_j p(\pi_j)$ (it is this $\max p(\pi_j)$ that is the issue and which makes the problem difficult, if we used $\min p(\pi_j)$ we could proceed as in the case $r = 1$). In applying the above lemma one loses all information in terms of specifying $p(\Gamma\backslash G)$. While the analysis can be made effective in principle, doing so would be unwieldy and the bound would anyway depend very poorly on Γ. For arithmetic applications the latter is a serious defect. We remark that in this case that G has no compact factors we don't know if $E(\Gamma\backslash G)$ is necessarily finite.

When G has a compact factor the situation is apparently more difficult. In the first place it is not known in general that $\Gamma\backslash G$ has a strong spectral gap. The most problematic case is the simplest one, that is $G = SL(2,\mathbb{R}) \times \mathrm{SU}(2)$. The suggestion (2) in [15, page 57] is equivalent to the existence of a strong spectral gap for any irreducible Γ in such a G. In [15] this spectral gap is proved for many Γ's and this has been extended (using novel methods from additive combinatorics) in [5] to include any Γ whose projection on SU(2) consists of matrices with algebraic numbers as entries. However in this case of G having compact factors, the set $E(\Gamma\backslash G)$ can be far more complicated. Borrowing a technique in [32] we show the following

Theorem 1. *There is an irreducible* Γ *in* $G = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SU}(2)$ *for which* $E(\Gamma\backslash G)$ *is infinite. In fact this is so for the generic* Γ.

Next, we turn to the simplest and most basic case for which an effective spectral gap is not known, that is for Γ an irreducible co-compact lattice in $G = \mathrm{PSL}(2,\mathbb{R})^d$, $d \geq 2$. Such a Γ is arithmetic and from the classification of such groups [48] we have that Γ is commensurable with the unit group in a suitable division algebra (see section 1.2). Serre conjectures that the congruence subgroup property holds for such groups, this being the most elementary and fundamental case for which the congruence subgroup problem is open (see [34, Chapter 7]). If true, this coupled with the Jacquet-Langlands correspondence [21] yields that $E(\Gamma\backslash G)$ is empty if the Ramanujan-Selberg conjecture [44] is true, and that $E(\Gamma\backslash G) \subseteq (2, \frac{18}{7}]$ using [24].

We can now formulate our main result. We will work in a slightly more general setting allowing Γ to act via a unitary representation. Let $\rho : \Gamma \to U(N)$ be an N-dimensional unitary representation of Γ. Let $L^2(\Gamma\backslash G, \rho)$ denote the space of functions from G to $\mathbb{C}^N$ satisfying

$$\begin{aligned} f(\gamma g) &= \rho(\gamma) f(g), \\ \int_{\Gamma\backslash G} &|f(g)|^2 dg < \infty. \end{aligned} \tag{0.3}$$

The regular representation $f(x) \mapsto f(xg)$ of G on $L^2(\Gamma\backslash G, \rho)$ decomposes discretely as

$$L^2(\Gamma\backslash G, \rho) \cong \bigoplus_{k=0}^{\infty} \pi_k(\rho), \tag{0.4}$$

with $\pi_k(\rho)$ irreducible representations of G.

Theorem 2. *Let* $\Gamma \subseteq \mathrm{PSL}(2,\mathbb{R})^d$ *be an irreducible co-compact lattice and* ρ *and* $\pi_k(\rho)$ *be as above. Then for any* $\alpha > 0$, $p(\pi_k(\rho)) < 6 + \alpha$ *except for at most a finite number of* k*'s. In particular*

$$E(\Gamma\backslash G) \cap [6 + \alpha, \infty)| < \infty.$$

Remark 0.1. From the arithmeticity of Γ $(n > 2)$, we know that it is commensurable to a lattice $\Gamma_{\mathcal{A}}$ derived from a quaternion algebra. We can thus assume (replacing Γ by $\Gamma \cap \Gamma_{\mathcal{A}}$ if necessary) that $\Gamma \subseteq \Gamma_{\mathcal{A}}$ is a finite index subgroup. Moreover, since we can also replace the representation ρ by the induced representation $\mathrm{Ind}_{\Gamma}^{\Gamma_{\mathcal{A}}}\rho$, it is sufficient to prove the theorem only in the case where $\Gamma = \Gamma_{\mathcal{A}}$.

Remark 0.2. The theorem implies that $p(\Gamma, \rho) < \infty$ and much more. The proof yields effective bounds (polynomial in $\dim \rho$) both for the number of exceptions as well as bounds for $p(\pi_k)$ for these exceptions. For some applications the finite number of exceptions enter as secondary terms in rates of equidistribution and are harmless, so that the theorem is effectively asserting that $p(\Gamma, \rho) \leq 6$.

Remark 0.3. The proof of the theorem is based on the Selberg trace formula [13, 16] and counting arguments involving relative quadratic extensions of L (the field of definition of the corresponding quaternion algebra) as in [46]. One can probably combine the analysis here with that in [41] (see also [20]) to show that for any fixed Γ as above and any Λ a congruence subgroup of Γ (i.e., the intersection of Γ with a congruence subgroup of the unit group of the quaternion algebra) that the exceptional π_k's for $L^2(\Lambda\backslash G)$ with $p(\pi_k) > 6 + \alpha$, consists only of the finite number of π_k's that are there from Γ (i.e., no new exceptional π's appear in passing from Γ to Λ). We have not carried this out and doing so would be of interest since for most applications this uniform spectral gap is a good substitute for the Ramanujan Conjectures.

We apply the theorem to the Selberg Zeta function in this setting. For simplicity we take $d = 2$ and Γ torsion free. Each $1 \neq \gamma \in \Gamma$ is of the form (γ_1, γ_2) with $\gamma_j \in \mathrm{PSL}(2, \mathbb{R})$ and $\gamma_j \neq 1$. We call γ mixed if γ_1 is hyperbolic and γ_2 is elliptic. That is γ_1 is conjugate to $\begin{pmatrix} N(\gamma)^{1/2} & 0 \\ 0 & N(\gamma)^{-1/2} \end{pmatrix}$ with $N(\gamma) > 1$ and γ_2 is conjugate to $\begin{pmatrix} \epsilon(\gamma) & 0 \\ 0 & \epsilon(\gamma) \end{pmatrix}$ with $|\epsilon(\gamma)| = 1$. For $m \geq 1$, Selberg [43] defines a Zeta function (see also [37])

$$Z_m(s, \Gamma) = \prod_{\{\gamma\}_\Gamma^*} \prod_{\substack{\nu=0 \\ |i|<m}}^{\infty} (1 - \epsilon_\gamma^i N(\gamma)^{-s-\nu})^{-1} \tag{0.5}$$

where the product is over all primitive conjugacy classes of mixed elements in Γ. He shows that $Z_m(s, \Gamma)$ is entire (except when $m = 1$ where it has a simple pole at $s = 1$) and satisfies a functional equation relating s and $1-s$. Its zeros are either real in $\{-k\}_{k>0} \cup (-1,1)$ or complex in $\frac{1}{2}+i\mathbb{R}$. They correspond to the eigenvalues of the Casimir operator acting on suitable functions on $\Gamma\backslash G$. As Selberg remarks, the form that these Zeta functions take is qualitatively similar to the Riemann Zeta function. In fact more so then the case of one upper half plane where the corresponding definition to (0.5) doesn't have a -1 in the exponent (this feature is connected with the parity of d). If Γ is a congruence group and the Ramanujan-Selberg conjecture is true then $Z_m(s,\Gamma)$ satisfies the "Riemann hypothesis", that is all its non trivial zeros are on $\Re(s)=\frac{1}{2}$.

As a corollary of Theorem 2 we get a zero free region that holds for all (but finitely many) of these Zeta functions.

Corollary 0.1. *Given $t_0 > \frac{5}{6}$ there is $m_0(\Gamma)$ such that $Z_m(s,\Gamma)$ has no zeros in $\Re(s) > t_0$, for $m \geq m_0(\Gamma)$.*

We now outline the main ideas of the proof of Theorem 2 (for the case $d=2$). As mentioned above it is sufficient to give a proof for Γ a lattice derived from a quaternion algebra, $\mathcal{A}$, defined over a number field L and an arbitrary unitary representation ρ of Γ. What we will actually show, is that if a representation $\pi \cong \pi_1 \otimes \pi_2$ occurs in the decomposition of $L^2(\Gamma\backslash G,\rho)$ with $p(\pi)$ sufficiently large, then all the spectral parameters of π are bounded.

We assume that $\pi \cong \pi_1 \otimes \pi_2$ occurs in the decomposition with say $p(\pi_1) > 6$ and the other spectral parameter large and get a contradiction: From our assumption $\pi_1 \cong \pi_{s_1}$ is complementary with $|s_1 - \frac{1}{2}| \in (\frac{1}{3},\frac{1}{2})$ and the second factor is either principal series $\pi_2 \cong \pi_{s_2}$ with $s_2 = \frac{1}{2}+ir_2,\ r_2 \in [T,2T]$, or discrete series $\pi_2 \cong \mathfrak{D}_m$ with weight $m \in [T,2T]$ for some large T. Let $g_1, g_2 \in C^\infty(\mathbb{R})$ be smooth even real valued compactly supported functions such that their Fourier transforms $h_j = \hat{g}_j$ are positive on $\mathbb{R}\cup i\mathbb{R}$. Further assume that h_2 vanishes at zero to a large order (for the discrete series case instead of h_2 we will use $\psi \in C^\infty(\mathbb{R})$ that is smooth, positive and compactly supported away from zero). For T large and $R = c\log(T)$ we have $h_1(Rr_1)h_2(\frac{r_2}{T}) \gg \frac{T^{c|1/2-s_1|}}{\log(T)}$ (equivalently in the second case the same bound holds for $h_1(Rr_1)\psi(\frac{m}{T})$). From the positivity assumption, this lower bound also holds when summing over all representation in the decomposition (in the second case we also sum over all weights $m \in \mathbb{Z}$). For the full sum we can also give an upper bound of order $O_\epsilon(T^2 + T^{c/2+\epsilon-1})$. For $c = 6-2\epsilon$ and T sufficiently large the upper bound is already smaller then the lower bound excluding the existence of such a representation in the decomposition.

Remark 0.4. When summing over all representation the trivial representation $r_0 = (i/2, i/2)$ also contributes. If h_2 (equivalently ψ) did not vanish at zero then the trivial representation would contribute $\sim T^{c/2}$ which is already larger then the lower bound coming from the representation we wish to exclude. Hence, in order for this strategy to have any chance of working we must make the function h_2 vanish at zero (or respectively take ψ supported away from zero).

To obtain the upper bound for the full sum we use the Selberg Trace formula to transform the spectral sum to a sum over the conjugacy classes (when summing over the weights we also use Poisson summation). We then bound each summand by its absolute value. (Even though the summands here are not positive, it turns out that the oscillations are sufficiently slow so that we apparently don't lose to much.) After some standard manipulation, using the fact that the test functions are compactly supported, estimating the sum over the conjugacy classes amounts to two counting arguments. The first is an estimate on the number of algebraic integers in L (viewed as a lattice in $\mathbb{R}^n$) that lie inside a long and narrow rectangular box whose sides are parallel to the coordinates axes. Using a simple Dirichlet box principle argument we bound the number of such lattice points by the volume of the box. The second counting problem is counting the number of conjugacy classes in Γ with a given trace, which amounts to estimating the number of optimal embeddings of certain orders into the quaternion algebra. This in turn is translated (via the work of Eichler) to estimates of class numbers of quadratic extensions of the number field L, that we obtain using Dirichlet's class number formula.

Acknowledgements. We thank A.Gamburd and T.Venkataramana for discussions about various aspects of the paper.

1. Background and Notation

In this section we go over some necessary background on lattices Γ in $G = \mathrm{PSL}(2, \mathbb{R})^d$, on the spectral decomposition of $L^2(\Gamma\backslash G)$ and the Selberg trace formula.

1.1. **Irreducible lattices.** A discrete subgroup $\Gamma \subset G = \mathrm{PSL}(2, \mathbb{R})^d$ is called a lattice if the quotient $\Gamma\backslash G$ has finite volume, and co-compact when $\Gamma\backslash G$ is compact. We say that a lattice $\Gamma \subset G$ is irreducible, if for every (non-central) normal subgroup $N \subset G$ the projection of Γ to G/N is dense. An equivalent condition for irreducibility is that for any nontrivial $1 \neq \gamma \in \Gamma$, none of the projections $\gamma_j \in G_j$ are trivial [46, Theorem 2]. Examples of irreducible lattices can be constructed from norm one elements of orders in a quaternion algebra (see below).

Recall that a nontrivial element $g \in \mathrm{PSL}(2,\mathbb{R})$ is called hyperbolic if $|\mathrm{Tr}(g)| > 2$, elliptic if $|\mathrm{Tr}(g)| < 2$, and parabolic if $|\mathrm{Tr}(g)| = 2$. For any nontrivial $1 \neq \gamma \in \Gamma$, the projections to the different factors are either hyperbolic or elliptic. The irreducibility implies that there are no trivial projections and since we assume Γ is co-compact there are no parabolic projections. There are purely hyperbolic elements (where all projections are hyperbolic), and mixed elements (where some projections are hyperbolic and other are elliptic). There could also be a finite number of torsion points that are purely elliptic (see e.g., [13, 46]).

1.2. **Lattices derived from quaternion algebras.** Let L be a totally real number field and denote by $\iota_1, \dots, \iota_n$ the different embeddings of L into $\mathbb{R}$. Let $\mathcal{A}$ be a quaternion algebra over L, ramified in all but d of the real places (say $\iota_1, \dots, \iota_d$). That is we have that $\mathcal{A} \otimes_{\iota_j(L)} \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$ for $j \leq d$ and it is isomorphic to the standard Hamilton's quaternions for $j > d$. Let $\mathcal{R}$ be a maximal order inside $\mathcal{A}$, and denote by $\mathcal{R}^1$ the group of (relative) norm one elements inside this order. The image $\iota_j(\mathcal{R}^1) \subset \mathrm{SL}(2,\mathbb{R})$ for $j \leq d$ and $\iota_j(\mathcal{R}^1) \subseteq SU(2)$ for $j > d$. The group $\Gamma(\mathcal{R}) = \left\{(\iota_1(\alpha), \dots, \iota_d(\alpha)) \in \mathrm{PSL}(2,\mathbb{R})^d | \alpha \in \mathcal{R}^1\right\}$ is a lattice inside $\mathrm{PSL}(2,\mathbb{R})^d$ and it is co-compact unless $n = d$ and $\mathcal{A} = \mathrm{Mat}(2, L)$ (see [45, 46]). Margulis's arithmeticity theorem [35, Chapter IX] together with Weil's classification of arithmetic lattices [48] implies that, up to commensurability, these are the only examples of irreducible co-compact lattices in $\mathrm{PSL}(2,\mathbb{R})^d$, $d \geq 2$.

1.3. **Spectral decomposition.** Let Γ be an irreducible co-compact lattice in G and let ρ be a finite dimensional unitary representation of Γ. The space $L^2(\Gamma\backslash G, \rho)$ is the space of Lebesgue measurable vector valued functions on G satisfying that $f(\gamma g) = \rho(\gamma) f(g)$ and that $\int_{\Gamma\backslash G} |f(g)|^2 dg \leq \infty$. The group G acts on $L^2(\Gamma\backslash G, \rho)$ by right multiplication and we can decompose it into irreducible representations

$$L^2(\Gamma\backslash G, \rho) \cong \bigoplus \pi_k.$$

Any irreducible unitary representation π_k is a product $\pi_k \cong \pi_{k,1} \otimes \pi_{k,2} \ldots \otimes \pi_{k,d}$ where the $\pi_{k,j}$'s are irreducible unitary representations of $\mathrm{PSL}(2,\mathbb{R})$. We briefly recall the classification of these representations. Other then the trivial representation the irreducible representations of $\mathrm{PSL}(2,\mathbb{R})$ are either principal series π_s, $s \in \frac{1}{2} + i\mathbb{R}$, complementary series π_s, $s \in (0,1)$, or discrete series $\mathfrak{D}_m$, $m \in \mathbb{Z}$. The discrete and principal series are both tempered, while the complementary series is non-tempered with $p(\pi_s) = \max\left\{\frac{1}{s}, \frac{1}{1-s}\right\}$. For a representation $\pi_k \cong \pi_{k,1} \otimes \pi_{k,2} \ldots \otimes \pi_{k,d}$ of G we have that $p(\pi_k) = \max_j p(\pi_{k,j})$.

1.4. **The Selberg Trace Formula.** The Selberg trace formula relates the spectral decomposition of $L^2(\Gamma\backslash G, \rho)$, to the conjugacy classes in Γ. We refer to [13, Sections 1-6], [16, Chapter 3] and [43] for the full derivation of the trace formula in this setting.

Fix a wight $m \in \mathbb{Z}^d$. For simplicity, we assume that $m_j = 0$ for $j \leq d_0$ and that $|m_j| > 1$ for $j > d_0$. Denote by $L^2(\Gamma\backslash G, \rho, m)$ the subspace of $L^2(\Gamma\backslash G, \rho)$ such that $\pi \cong \pi_1 \otimes \pi_2 \cdots \otimes \pi_d$ occurs in the decomposition if and only if π_j is principal or complementary series for $j \leq d_0$ and $\pi_j \cong \mathfrak{D}_{m_j}$ for $j > d_0$. Consider the decomposition

$$L^2(\Gamma\backslash G, \rho, m) \cong \bigoplus_{k=0}^{\infty} \pi_k,$$

into irreducible representations. For any $j \leq d_0$ let $s_{k,j} = \frac{1}{2} + ir_{k,j}$ such that $\pi_{k,j} = \pi_{s_{k,j}}$. For any $j \leq d_0$ let $g_j \in C^\infty(\mathbb{R})$ be a smooth even real valued compactly supported function, and let $h_j = \hat{g}_j$ be its Fourier transform. Recall that for any $\gamma \in \Gamma$ its projections to the different factors are either hyperbolic, $\gamma_j \sim \begin{pmatrix} N(\gamma_j)^{1/2} & 0 \\ 0 & N(\gamma_j)^{-1/2} \end{pmatrix}$ with $N(\gamma_j) = e^{l_j} > 1$, or elliptic $\gamma_j \sim \begin{pmatrix} \epsilon(\gamma_j) & 0 \\ 0 & \epsilon(\gamma_j) \end{pmatrix}$ with $\epsilon(\gamma_j) = e^{i\theta_j} \in S^1$. Define the functions $\tilde{h}_j(\gamma_j)$ by

$$\tilde{h}_j(\gamma_j) = \frac{g(l_j)}{\sinh(l_j/2)},$$

when γ_j is hyperbolic, and

$$\tilde{h}_j(\gamma_j) = \frac{1}{\sin\theta_j} \int_{-\infty}^{\infty} \frac{\cosh[(\pi - 2\theta_j)r]}{\cosh(\pi r)} h(r)dr$$

when γ_j is elliptic. The Selberg trace formula, applied to the product $h(r) = \prod_{j \leq d_0} h_j(r_j)$, then takes the form

$$\sum_k h(r_k) =$$

$$\frac{\mathrm{vol}(\Gamma\backslash G)\chi_\rho(1)}{(4\pi)^d} \prod_{j \leq d_0} \left(\int_{\mathbb{R}} h_j(r_j) r_j \tanh(\pi r_j) dr_j \right) \prod_{j > d_0} (2|m_j| - 1)$$

$$+ \sum_{\{\gamma\}} \mathrm{vol}(\Gamma_\gamma \backslash G_\gamma) \chi_\rho(\gamma) \prod_{j \leq d_0} \tilde{h}_j(\gamma_j) \prod_{j > d_0} \frac{e^{\pm 2i|m_j|\theta_j}}{1 - e^{\pm 2i\theta_j}},$$

where the sum on the right hand side is over all Γ-conjugacy classes $\{\gamma\} \in \Gamma^\sharp$ that are elliptic for $j > d_0$, where G_γ denotes the centralizer of γ in G and $\Gamma_\gamma = G_\gamma \cap \Gamma$, $\chi_\rho(\gamma) = \mathrm{Tr}(\rho(\gamma))$ is the character of the representation, and the $\pm$ signs are determined by the signs of the m_j's.

2. Proof of Theorem 1

In this section we give the proof of Theorem 1. We consider $\mathrm{SU}(2) \times \mathrm{SU}(2)$ as a deformation space for lattices Γ in $G = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SU}(2)$. We construct a dense set of irreducible lattices inside this deformation space, each satisfying that $E(\Gamma\backslash G)$ is infinite, and then use these to show that the same is true generically.

2.1. **Deformation space.** Let Γ be an irreducible lattice in $G = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SU}(2)$. The projection P_1 of Γ onto the first factor has image Λ which is a lattice in $\mathrm{SL}(2,\mathbb{R})$. For the purpose of constructing lattices Γ in G with $|E(\Gamma\backslash G)| = \infty$, we assume that $P_1 : \Gamma \to \Lambda$ is an isomorphism. In this way we can identify

$$\Gamma = \{(\gamma, \rho(\gamma)) : \gamma \in \Lambda\},$$

where $\rho = P_1 \circ P_2^{-1} : \Lambda \hookrightarrow \mathrm{SU}(2)$. For Λ we take the congruence subgroup $\Gamma(2)$ of $\mathrm{SL}(2,\mathbb{R})$ which is a free group on two generators $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. Our deformation space of such lattices can then be described as $\mathrm{SU}(2) \times \mathrm{SU}(2)$ where for any $u = (u_1, u_2) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$ define ρ_u by $\rho_u(A) = u_1,\ \rho_u(B) = u_2$ extended to a homomorphism of Λ into $\mathrm{SU}(2)$ and let $\Gamma_u = (\Lambda, \rho_u)$. One can further identify such lattices in G which are conjugate in G but for our analysis there is no need to do so. For any $n \geq 3$ we choose u_1, u_2 so that $u_1^n = u_2^n = 1$ and satisfy no further relations (that is the corresponding image $\rho_u(\Lambda)$ is isomorphic to the free product $(\mathbb{Z}/n\mathbb{Z}) * (\mathbb{Z}/n\mathbb{Z})$). Varying over all such ρ_u and all $n > 3$ yields a dense subset in our deformation space. Note that for any such choice of u the image $\rho_u(\Lambda)$ is dense in $\mathrm{SU}(2)$, that is, $\Gamma_u = (\Lambda, \rho_u)$ is irreducible. We will now show that for such a lattice we have $|E(\Gamma_u\backslash G)| = \infty$.

Theorem 3. *For any homomorphism $\rho_u : \Lambda \to \mathrm{SU}(2)$ as above the corresponding lattice $\Gamma_u = (\Lambda, \rho_u)$ satisfies $|E(\Gamma_u\backslash G)| = \infty$.*

2.2. **Spectral theory for infinite volume quotients of $\mathbb{H}$.** For the proof of Theorem 3 we will make a reduction to the spectral theory of $L^2(L\backslash\mathbb{H})$ with $\mathbb{H}$ the upper half plane and $L = \ker \rho_u$ acting by linear fractional transformations. Before proceeding with the proof we review some facts on the spectral theory of these infinite volume hyperbolic surfaces that we will need (we refer to [47] for details). Let L be a torsion free discrete subgroup of $\mathrm{SL}(2,\mathbb{R})$. Then $L\backslash\mathbb{H}$ is a complete hyperbolic surface and the Laplacian on smooth functions of compact support on $L\backslash\mathbb{H}$ has a unique self adjoint extension denoted by $\triangle$. Let $\lambda_0(L\backslash\mathbb{H})$ denote the bottom of the spectrum of $\triangle$ so that the spectrum is contained in $[\lambda_0,\infty)$. Closely related to $\lambda_0(L\backslash\mathbb{H})$ is the exponent of convergence $\delta(L)\in[0,1]$ (see [47, page 333] for definition). When L is nonelementary and contains a parabolic element this exponent $\delta(L)>\frac{1}{2}$ [2, Theorem 7], in which case the Elstrodt-Patterson Theorem [47, Theorem 2.17] says that $\lambda_0(L\backslash\mathbb{H})=\delta(L)(1-\delta(L))$, and in particular $\lambda_0(L\backslash\mathbb{H})<1/4$.

We shall be interested in the case where L is a normal subgroup of Λ and Λ/L is not amenable. In this case Brooks [6] shows that $\lambda_0(L\backslash\mathbb{H})>0$. Summarizing the above remarks we have

Proposition 2.1. *Let L satisfy that Λ/L is not amenable and $\delta(L)>1/2$ then $0<\lambda_0(L\backslash\mathbb{H})<1/4$ and λ_0 is an accumulation point of distinct points of the spectrum of $\triangle$ on $L^2(L\backslash\mathbb{H})$.*

Proof. From the above remarks it is clear that $0<\lambda_0(L\backslash\mathbb{H})<1/4$. We will show that there is no eigenfunction in $L^2(L\backslash\mathbb{H})$ with eigenvalue λ_0 implying that λ_0 cannot be an isolated point in the spectrum. We recall that if an eigenfunction $\phi\in L^2(L\backslash\mathbb{H})$ with eigenvalue λ_0 exists then it is unique up to a nonzero scalar multiple [47, Corollary 2.9]. On the other hand, as L is normal in Λ, for any $\gamma\in\Lambda$ the function $\phi(\gamma z)\in L^2(L\backslash\mathbb{H})$ is also a λ_0-eigenfunction. Consequently, we must have $\phi(z)=\phi(\gamma z)$ for all $\gamma\in\Lambda$ and since Λ/L is infinite then ϕ can not be in $L^2(L\backslash\mathbb{H})$. $\square$

Remark 2.1. The situation here is very different from the case of geometrically finite quotients where Lax and Phillips [27] showed that the point spectrum is finite. Indeed, we recall that a finitely generated normal subgroup of a free group is always of finite index [22]. Hence, the assumption that L is a normal subgroup with infinite index in Λ implies that L must be infinitely generated and in particular not geometrically finite.

2.3. **Construction of nontempered points.** Fix $n>3$ and a homomorphism $\rho:\Lambda\to\mathrm{SU}(2)$ with $\rho(\Lambda)\cong(\mathbb{Z}/n\mathbb{Z})*(\mathbb{Z}/n\mathbb{Z})$ such that $\rho(A)^n=\rho(B)^n=1$. The kernel $L=\ker(\rho)$ is normal in Λ and $\Lambda/L\cong(\mathbb{Z}/n\mathbb{Z})*(\mathbb{Z}/n\mathbb{Z})$ is infinite (and not amenable). Also $A^n\in L$ is parabolic so $\delta(L)>\frac{1}{2}$ and hence $0<\lambda_0(L\backslash\mathbb{H})<\frac{1}{4}$ is an accumulation point of distinct points in the spectrum.

Now, for $l \geq 0$ let $\sigma_l = \mathrm{sym}^l$ denote the $l+1$ dimensional irreducible representation of $\mathrm{SU}(2)$. According to Weyl $L^2(\mathrm{SU}(2))$ decomposes under the regular representation as

$$L^2(\mathrm{SU}(2)) = \bigoplus_{l=0}^{\infty} (\dim \sigma_l) W_l,$$

where $W_l \cong \sigma_l$. Correspondingly the regular representation of G on $L^2(\Gamma \backslash G)$ decomposes into the representations $L^2(\Lambda \backslash \mathrm{SL}(2,\mathbb{R}), \sigma_l \circ \rho)$ each occurring with multiplicity $l+1$. Here

$$\begin{aligned} (2.1)\quad & L^2(\Lambda \backslash \mathrm{SL}(2,\mathbb{R}), \sigma_l \circ \rho) = \\ & \qquad \left\{ F : \mathrm{SL}(2,\mathbb{R}) \to \mathbb{C}^{l+1} | F(\gamma g) = \sigma_l(\rho(\gamma)) F(g),\ \gamma \in \Lambda \right\}, \end{aligned}$$

with the right action of $\mathrm{SL}(2,\mathbb{R})$ (we may normalize so that σ_l acts unitarily on $\mathbb{C}^{l+1}$ with respect to the standard inner product). Since we are only interested in representations π of $\mathrm{SL}(2,\mathbb{R})$ appearing in (2.1) which are nontempered, we may restrict to π's which are spherical.

Denote by $L^2(\Lambda \backslash \mathbb{H}, \sigma_l \circ \rho)$ the space of square integrable vector valued functions on the upper half plane $\mathbb{H}$, satisfying $F(\gamma z) = \sigma_l(\rho(\gamma))F(z)$, where γ acts on $z \in \mathbb{H}$ by fractional linear transformations. This space is naturally identified with the space of spherical vectors in $L^2(\Lambda \backslash \mathrm{SL}(2,\mathbb{R}), \sigma_l \circ \rho)$. Let $F_{1,l}, F_{2,l}, \ldots$ in $L^2(\Lambda \backslash \mathbb{H}, \sigma_l \circ \rho)$ be an orthonormal basis of eigenvectors of $\triangle$ with eigenvalues $\lambda_{j,l} = \frac{1}{4} + t_{j,l}^2$ giving the discrete spectrum and $E(z, \frac{1}{2} + it)$, $t \in \mathbb{R}$ spanning the (tempered) continuous spectrum. Note that if $\triangle F_{j,l} + \lambda_{j,l} F_{j,l} = 0$ with $\lambda_{j,l} = s_{j,l}(1 - s_{j,l}) < \frac{1}{4}$ then there is a nontempered representation π appearing in $L^2(\Gamma \backslash G)$ with $p(\pi) = \frac{1}{1 - s_{j,l}}$. Hence showing that $|E(\Gamma \backslash G)| = \infty$ is equivalent to showing that there are infinitely many distinct eigenvalues $\lambda_{j,l}$ below $1/4$. The following proposition then concludes the proof of Theorem 3.

Proposition 2.2. *With the above notations, there are infinitely many eigenvalues $\lambda_0(L \backslash \mathbb{H}) < \lambda_{j,l} < \frac{1}{4}$ accumulating at $\lambda_0(L \backslash \mathbb{H})$.*

Proof. Let $k(z,w)$ be a point pair invariant on $\mathbb{H}$ as in [44] (i.e., for any $g \in \mathrm{SL}(2,\mathbb{R})$, $k(gz, gw) = k(z,w)$). We assume that for z fixed $k(z,w)$ is a continuous compactly supported function in w. We have the spectral expansion for the kernel $K_{\sigma_l}(z,w)$ (see [17, Chapter 8, equation 4.1]) given by

$$(2.2)\qquad \begin{aligned} K_{\sigma_l}(z,w) &= \sum_{\gamma\in\Lambda} k(\gamma z,w)\sigma_l\circ\rho(\gamma) \\ &= \sum_{j=1}^{\infty} h(t_{j,l})F_{j,l}(z)\overline{F^t_{j,l}(w)} \\ &+ \int_{\mathbb{R}} h(t)E(z,\frac{1}{2}+it)\overline{E^t(w,\frac{1}{2}+it)}dt, \end{aligned}$$

where $h(s)=\int_{\mathbb{H}} k(i,z)y^s\frac{dxdy}{y^2}$ is the Selberg transform of k. Note that for any fixed z,w both sides are $(l+1)\times(l+1)$ matrices. Taking traces of these matrices gives

$$(2.3)\qquad \begin{aligned} \sum_{\gamma\in\Lambda} k(\gamma z,w)\chi_l(\rho(\gamma)) &= \sum_{j=1}^{\infty} h(t_{j,l})\langle F_{j,l}(z),F_{j,l}(w)\rangle \\ &+ \int_{\mathbb{R}} h(t)\langle E(z,\frac{1}{2}+it),E(w,\frac{1}{2}+it)\rangle dt, \end{aligned}$$

where χ_l is the character of σ_l on SU(2) and we denote by $\langle,\rangle$ the standard inner product on $\mathbb{C}^{l+1}$.

Let $\psi(z)$ be a continuous function of compact support in $\mathbb{H}$ and integrate (2.3) against $\psi(z)\overline{\psi(w)}$ to get

$$(2.4)\qquad \begin{aligned} \frac{1}{l+1}\sum_{\gamma\in\Lambda}\int_{\mathbb{H}}\int_{\mathbb{H}}\psi(z)\overline{\psi(w)}k(\gamma z,w)dv(z)dv(w)\chi_l(\rho_n(\gamma)) \\ = \int_C h(t)d\mu_l(t) \end{aligned}$$

where μ_l is the positive measure on $C=[0,\infty)\cup[0,\frac{i}{2}]$ given by

$$(2.5)\qquad \begin{aligned} &\frac{1}{l+1}\sum_{j=1}^{\infty}\langle\int_{\mathbb{H}}\psi(z)F_{j,l}(z)dv(z),\int_{\mathbb{H}}\psi(z)F_{j,l}(w)dv(w)\rangle\delta_{t_{j,l}} \\ &+\frac{1}{l+1}\langle\int_{\mathbb{H}}\psi(z)E(z,\frac{1}{2}+it)dv(z),\int_{\mathbb{H}}\psi(w)E(w,\frac{1}{2}+it)dv(w)\rangle dt \end{aligned}$$

Note that for fixed k the sum over Λ on the left hand side of 2.4 is finite. Also as $l\to\infty$ we have $\frac{1}{l+1}\chi_l(u)\to 1$ if $u=1$ and tends to 0 if $u\neq 1$. Hence, taking the limit $l\to\infty$ in (2.4) (for k and ψ fixed) we get that

$$(2.6)\qquad \mu_l(h)\to\sum_{\gamma\in L}\int_{\mathbb{H}}\int_{\mathbb{H}}\psi(z)\overline{\psi(w)}k(\gamma z,w)dv(z)dv(w).$$

If the function $\psi(z)$ is supported in a small ball B in $\mathbb{H}$ that is contained in one fundamental domain $\mathcal{F}$ for $L\backslash\mathbb{H}$ then we can think of ψ also as an element of $L^2(L\backslash\mathbb{H})$. For such ψ we get

$$\mu_l(h) \to \int_{\mathcal{F}}\int_{\mathcal{F}} \psi(z)\overline{\psi(w)}K_L(z,w)dv(z)dv(w), \tag{2.7}$$

where

$$K_L(z,w) = \sum_{\gamma\in L} k(\gamma z, w). \tag{2.8}$$

The function $K_L(z,w)$ is $L\times L$ invariant and gives a kernel for a bounded self-adjoint operator on $L^2(L\backslash\mathbb{H})$. The family of such operators (when taking different point pair invariants k) is a commutative algebra that also commutes with $\triangle$. Consequently, this whole algebra can be simultaneously diagonalized together with $\triangle$. For any fixed $\psi\in L^2(L\backslash\mathbb{H})$ there is a corresponding positive spectral measure ν_ψ on the spectrum of $\triangle$. That is, using the parameter $t=\sqrt{\lambda-1/4}$ we have the spectral decomposition

$$\langle K_L\psi,\psi\rangle = \int_C h(t)d\nu_\psi(t). \tag{2.9}$$

Consequently, from (2.7) and (2.9) we get that for every function h which is the Selberg transform of k continuous of compact support (in particular for any even function h with Fourier transform smooth of compact support) as $l\to\infty$

$$\mu_l(h)\to\nu_\psi(h). \tag{2.10}$$

Now, since the spectrum of $\triangle$ on $L^2(L\backslash\mathbb{H})$ has λ_0 as an accumulation point it follows that given $\epsilon>0$ we can find a closed nonempty subinterval I of $(\lambda_0,\lambda_0+\epsilon)$ such that the spectral projector P_I onto I is nonzero. Let f be a nonzero element in the image of this projector P_I. One can choose a small ball B in $\mathbb{H}$ which is injective in $L\backslash\mathbb{H}$ and such that f restricted to B is a nonzero L^2 function. Take ψ to be supported in B, continuous and such that its integral over B against f is not zero. Then as members in $L^2(L\backslash\mathbb{H})$ the inner product of f and ψ is not zero so that the support of ν_ψ meets I nontrivially. Let $J\subset(\lambda_0,\lambda_0+\epsilon)$ be an interval strictly containing I and let h be an even function with Fourier transform compactly supported that is negative outside J and satisfies that $\nu_\psi(h)>0$. Then from (2.10) (with this ψ and h) we get that for sufficiently large l the support of μ_l in (2.5) meets J nontrivially. Consequently, for all sufficiently large l there is an eigenvalue $\lambda_{j,l}\in(\lambda_0,\lambda_0+\epsilon)$. Repeating this procedure (making ϵ smaller) will produce infinitely many eigenvalues accumulating at λ_0.

To conclude the proof we give a construction for an even function h with Fourier transform smooth and compactly supported that is negative outside J and satisfies that $\nu_\psi(h) > 0$. Fix a smooth compactly supported function g with Fourier transform $\hat{g}$ even and positive on C and set $M > \frac{\int_C \hat{g}(t)d\nu_\psi(t)}{\int_I \hat{g}(t)d\nu_\psi(t)}$ (this is finite since the support of ν_ψ meets I). Now let $F(t) = \sum_{n\le N} a_n \cos(nt)$, be a trigonometric polynomial satisfying that $F(t) > M$ for $t^2 + \frac{1}{4} \in I$ and $-1 < F(t) < 0$ for $t^2 + \frac{1}{4}$ in the complement of J. (The existence of such a trigonometric polynomial is guaranteed by the Weierstrass's approximation theorem for polynomials recalling that the Chebyshev polynomials satisfy $T_n(\cos(t)) = \cos(nt)$). Now the function $h(t) = \hat{g}(t)F(t)$ has Fourier transform smooth of compact support and satisfies $h(t) < 0$ on the complement of J (as it has the same sign as F) and $\nu_\psi(h) > 0$ (by the choice of M). □

We now complete the proof of Theorem 1, showing that for generic u the exceptional exponent set $E(\Gamma_u\backslash G)$ is infinite. As we noted and is easily shown, the set of u's that we consider in Theorem 3 are dense in $\mathrm{SU}(2)\times\mathrm{SU}(2)$. Let $u_j,\ j = 1, 2, \dots$ be an enumeration of a dense set of such u's. Now for each l, the spectrum in $[0, 1/4]$ of $\triangle$ on $L^2(\Lambda\backslash\mathbb{H}, \sigma_l\circ\rho)$ is continuous in u. Hence it follows from Theorem 3 that for each $j = 1, 2, 3, \dots$ there is ϵ_j such that for u in a small neighborhood $B(u_j, \epsilon_j)$ of u_j the lattice $\Gamma_u = (\Lambda, \rho_u)$ satisfies $|E(\Gamma_u\backslash G)| > j$. Now let

$$B = \bigcap_{J=1}^{\infty} \bigcup_{j=J}^{\infty} B(u_j, \epsilon_j).$$

Then B is of the second category in $\mathrm{SU}(2)\times\mathrm{SU}(2)$ and for any $u \in B$, $E(\Gamma_u\backslash G)$ is infinite. We have thus shown that a generic lattice in the sense of Baire has infinitely many exceptional exponents. Note that for the generic $u \in \mathrm{SU}(2)\times\mathrm{SU}(2)$, u_1 and u_2 generate a free group in $\mathrm{SU}(2)$. Hence the limit measure in (2.10) (as $l \to \infty$) for such a lattice is supported on $\mathbb{R}$ (i.e., it has no exceptional spectrum). That is the generic lattice has infinitely many exceptional exponents but in terms of density almost all the representations are tempered.

3. Proof of Theorem 2

We now give the proof of Theorem 2. In order to simplify notations we will write down the full details only for the case $d = 2$. The modifications required to handle $d > 2$ are straight forward and are accounted for in section 3.3.

3.1. **Reduction to an asymptotic argument.** Fix a co-compact irreducible lattice, $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})^2$, derived from a quaternion algebra and let ρ be a unitary representation of Γ. It is well known that there are only finitely many representation occurring in $L^2(\Gamma\backslash G,\rho)$ with all spectral parameters bounded. We can thus reduce Theorem 2 to the following asymptotic argument

Theorem 4. *Assume that $\pi \cong \pi_1 \otimes \pi_2$ occurs in $L^2(\Gamma\backslash G,\rho)$ and that $\pi_1 \cong \pi_{s_1}$ is complementary series and π_2 is either principal π_{s_2} with $s_2 = \frac{1}{2} + ir_2,\ r_2 \in [T,2T]$ or discrete $\mathfrak{D}_m$ with $|m| \in [T,2T]$. Then for any $c > 0$ (as $T \to \infty$)*

$$T^{c|\frac{1}{2}-s_1|} \ll_\epsilon \dim(\rho)(\frac{T^2}{\log(T)} + T^{\frac{c}{2}-1+\epsilon})$$

We now show that this asymptotic argument implies Theorem 2.

Proof of Theorem 2. Fix $\alpha,\epsilon > 0$ and let $M = M(\alpha,\epsilon) = \frac{6+\alpha}{\alpha-\epsilon(4+\alpha)}$. Then by theorem 4 with $c = 6-2\epsilon$, there is a constant $C = C(\epsilon,\Gamma)$ such that if $\pi = \pi_1 \otimes \pi_2$ occurs in the decomposition with π_1 complementary with $p(\pi_1) \geq 6+\alpha$ (i.e., $|\frac{1}{2}-s_1| \geq \frac{4+\alpha}{2(6+\alpha)}$) then π_2 is either complementary, or principal with parameter $r_2 \leq (C\dim\rho)^M$ or discrete with parameter $|m| \leq (C\dim\rho)^M$. Theorem 2 now follows as there are at most $O((\dim\rho)^{2M})$ such representations. $\square$

3.2. **Reduction to a counting argument.** We now use the Selberg trace formula to reduce Theorem 4 to a counting argument.

Proposition 3.1. *Assume that $\pi \cong \pi_1 \otimes \pi_2$ occurs in $L^2(\Gamma\backslash G,\rho)$ and satisfies the hypothesis of Theorem 4. Then for any $c > 0$ as $T \to \infty$*

$$T^{c|\frac{1}{2}-s_1|} \ll_\epsilon \dim(\rho)(\frac{T^2}{\log(T)} + T \sum_{\substack{|t_1|\leq T^{c/2} \\ |t_2|=2+O(T^{-2+\epsilon})}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}}$$
$$+\frac{1}{T}\sum_{\substack{|t_1|\leq T^{c/2} \\ |t_2|\leq 2}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}}).$$

where the summation is over elements $t = (t_1,t_2) \in Tr(\Gamma)$ and

$$F_\Gamma(t) = \sum_{\substack{\{\gamma\} \\ Tr(\gamma)=t}} \mathrm{vol}(\Gamma_\gamma\backslash G_\gamma),$$

is counting the number of conjugacy classes in Γ with a given trace.

We will give the proof separately for the two cases when π_2 is principal or discrete series.

Proof for principal series. Let $\pi \cong \pi_1 \otimes \pi_2$ occur in $L^2(\Gamma\backslash G, \rho)$ with $\pi_1 \cong \pi_{s_1}$ complementary series and $\pi_2 \cong \pi_{s_2}$ principal with parameter $r_2 \in [T, 2T]$. Consider the function

$$h_{R,T}(r_1, r_2) = h_1(Rr_1)h_2(\frac{r_2}{T}),$$

where h_1, h_2 are even positive functions with Fourier transforms g_1, g_2 smooth and supported on $[-1, 1]$. We also assume that h_2 vanishes at zero to a large order $> 2/\epsilon$. We note that this vanishing assumption is crucial for the proof (see remark 0.4). Note that for $s_1 = \frac{1}{2}+ir_1 \in (0, \frac{1}{2})$ and $r_2 \in [T, 2T]$ we can bound the function $h_{R,T}(r) \gg \frac{\exp(R|\frac{1}{2}-s_1|)}{R}$ from below. Since the function is positive, this is also a lower bound for the sum over the full spectrum

$$\frac{\exp(R|\frac{1}{2} - s_1|)}{R} \ll \sum_k h_{R,T}(r_k).$$

We now use the trace formula (with wight $(0, 0)$) to transform the sum over the eigenvalues to a sum over conjugacy classes. The geometric side of the trace formula is given by

$$\frac{\mathrm{vol}(\Gamma\backslash G)\chi_\rho(1)}{16\pi^2}\iint_{\mathbb{R}^2} h_{R,T}(r_1, r_2)r_1 \tanh(\pi r_1)r_2 \tanh(\pi r_2)dr_1dr_2$$

$$+\frac{T}{R}\sum_{\{\gamma\}\in e.h}\mathrm{vol}(\Gamma_\gamma\backslash G_\gamma)\frac{\chi_\rho(\gamma)g_2(Tl_{\gamma_2})}{\sinh(\frac{l_{\gamma_2}}{2})\sin\theta_{\gamma_1}}\int_{-\infty}^{\infty}\frac{\cosh(\frac{(\pi-2\theta_{\gamma_1})r}{R})}{\cosh(\frac{\pi r}{R})}h_1(r)dr$$

$$+\frac{T}{R}\sum_{\{\gamma\}\in h.e}\mathrm{vol}(\Gamma_\gamma\backslash G_\gamma)\frac{\chi_\rho(\gamma)g_1(\frac{l_{\gamma_1}}{R})}{\sinh(\frac{l_{\gamma_1}}{2})\sin\theta_{\gamma_2}}\int_{-\infty}^{\infty}\frac{\cosh[(\pi-2\theta_{\gamma_2})Tr]}{\cosh(\pi Tr)}h_2(r)dr$$

$$+\frac{T}{R}\sum_{\{\gamma\}\in h.h}\mathrm{vol}(\Gamma_\gamma\backslash G_\gamma)\frac{\chi_\rho(\gamma)g_1(\frac{l_{\gamma_1}}{R})}{\sinh(\frac{l_{\gamma_2}}{2})}\frac{g_2(Tl_{\gamma_1})}{\sinh(\frac{l_{\gamma_2}}{2})}$$

where we divided the conjugacy classes into the different types: Trivial conjugacy class, elliptic-hyperbolic, hyperbolic-elliptic and hyperbolic-hyperbolic. (There could also be elliptic-elliptic elements that we ignore as their total contribution to the sum is bounded by $O(1)$.) We will now give separate bounds for each term where we replace each summand by its absolute value and bound the character of the representation $|\chi_\rho(\gamma)| \leq \chi_\rho(1) = \dim\rho$ by the dimension.

Trivial Conjugacy class. By making a change of variables $r_1 \mapsto \frac{r_1}{R}$ and $r_2 \mapsto Tr_2$ and bounding $|\tanh(t)| \leq 1$, the contribution of the trivial conjugacy class is bounded by $O(\chi_\rho(1)\frac{T^2}{R^2})$.

Elliptic-Hyperbolic. For the elliptic-hyperbolic conjugacy class, note that g_2 is supported on $[-1,1]$, hence the only conjugacy classes contributing to this sum are the ones with $l_{\gamma_2} \leq \frac{1}{T}$. But there are only finitely many conjugacy classes with γ_1 elliptic and $l_{\gamma_2} \leq \frac{1}{T}$, hence the contribution of these conjugacy classes is bounded by $O(T)$. (In fact for T sufficiently large there are no conjugacy classes satisfying this condition so that it is bounded by $O(1)$).

Hyperbolic-Hyperbolic. The only contribution of hyperbolic-hyperbolic elements comes from elements with $l_{\gamma_1} \leq R$ and $l_{\gamma_2} \leq \frac{1}{T}$. It is convenient to rephrase this in terms of the traces of the conjugacy classes. For each conjugacy class, $\{\gamma\}$ it's trace $t = (t_1, t_2) = (\mathrm{Tr}(\gamma_1), \mathrm{Tr}(\gamma_2))$ is given by $t_j = e^{l_{\gamma_j}/2} + e^{-l_{\gamma_j}/2}$. Consequently, the only contribution comes from conjugacy classes such that $|t_1| \sim e^{l_{\gamma_1}/2} \leq e^{R/2}$ and $2 < |t_2| \leq 2 + \frac{1}{T^2}$. We can also write $\sinh(l_{\gamma_j}/2) = \sqrt{t_j^2 - 4}$ so the contribution of the hyperbolic-hyperbolic conjugacy classes is bounded by

$$(h.h.) \ll \dim(\rho)\frac{T}{R} \sum_{\substack{|t_1| \leq e^{R/2} \\ 2 < |t_2| < 2 + \frac{1}{T^2}}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2 - 4)(t_2^2 - 4)}}.$$

Hyperbolic-Elliptic. As above, since g_1 is supported on $[-1,1]$ the only contribution here is from conjugacy classes satisfying $l_{\gamma_1} \leq R$. For these we estimate the contribution of the integral

$$\int_{-\infty}^{\infty} \frac{\cosh[(\pi - 2\theta)Tr]}{\cosh(\pi Tr)} h_2(r) dr.$$

First for $\theta = \theta_{\gamma_2} < T^{-1+\epsilon}$ we just bound this integral by $O(1)$. Next, for $\theta > T^{-1+\epsilon}$ separate this integral into two parts: The first when r is small, where we just bound $|\frac{\cosh[(\pi-2\theta)Tr]}{\cosh(\pi Tr)}| \leq 1$ to get

$$|\int_{|r| \leq T^{-\epsilon/2}} \frac{\cosh[(\pi - 2\theta)Tr]}{\cosh(\pi Tr)} h_2(r) dr| \leq \int_{|r| \leq T^{-\epsilon/2}} h_2(r) dr.$$

Since we assume h_2 vanishes at zero to order$> \frac{2}{\epsilon}$ we get that $h_2(r) \ll r^{\frac{2}{\epsilon}}$ near zero, hence, the contribution of this part is bounded by $O(T^{-2})$. Now for the next part we can use the exponential decay of $\frac{\cosh[(\pi-2\theta)Tr]}{\cosh(\pi Tr)}$

to get that $\int_{|r|>T^{-\epsilon/2}} \frac{\cosh[(\pi-2\theta)Tr]}{\cosh(\pi Tr)} h_2(r)dr \ll e^{-T^{\epsilon/2}}$, so that for large T the whole integral is bounded by $O(\frac{1}{T^2})$.

Thus, for $|\theta_{\gamma_2}| < T^{-1+\epsilon}$ and $l_{\gamma_1} \le R$ (equivalently $2-\frac{1}{T^{2-\epsilon}} \le |t_2| \le 2$ and $|t_1| \le e^{R/2}$) we get a contribution of

$$\frac{T}{R}\frac{c_\gamma|\chi_\rho(\gamma)|}{\sinh(l_{\gamma_1}/2)\sin\theta_{\gamma_2}} = \frac{T}{R}\frac{c_\gamma|\chi_\rho(\gamma)|}{\sqrt{(t_1^2-4)|t_2^2-4|}},$$

and for $|\theta_{\gamma_1}| > T^{-1+\epsilon}$ and $l_{\gamma_2} \le R$ (equivalently $2-\frac{1}{T^{2-\epsilon}} \ge |t_2| \le 2$ and $|t_1| \le e^{R/2}$) we get a contribution of

$$\frac{1}{RT}\frac{c_\gamma|\chi_\rho(\gamma)|}{\sinh(l_{\gamma_1}/2)\sin\theta_{\gamma_2}} = \frac{1}{RT}\frac{c_\gamma|\chi_\rho(\gamma)|}{\sqrt{(t_1^2-4)|t_2^2-4|}}.$$

We can thus bound the contribution of the hyperbolic-elliptic elements by

$$\begin{aligned}(h.e.) \quad &\ll \quad \dim(\rho)\frac{T}{R}\sum_{\substack{|t_1|\le e^{R/2} \\ 2-\frac{1}{T^{2-\epsilon}}<|t_2|<2}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}}. \\ &+ \quad \dim(\rho)\frac{1}{RT}\sum_{\substack{|t_1|\le e^{R/2} \\ |t_2|\le 2}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}}\end{aligned}$$

Putting all these bounds together, and taking $R = c\log(T)$ concludes the proof. $\square$

Proof for discrete series. Let $\pi \cong \pi_1 \otimes \pi_2$ occur in $L^2(\Gamma\backslash G, \rho)$ with $\pi_1 \cong \pi_s$ complementary series with $s = \frac{1}{2} + ir \in (0, \frac{1}{2})$ and $\pi_2 \cong \mathfrak{D}_m$ discrete series with weight $m \in [T, 2T]$ (the case of $-m \in [T, 2T]$ is analogous). Let h be an even positive function satisfying $h(0) = 1$ with Fourier transforms g smooth and supported on $[-1, 1]$. Similar to the previous case, we can bound the function $h(Rr) \gg \frac{e^{R(1/2-s)}}{R}$ from below, and from positivity this is also a lower bound for the sum over all representations $\pi_k \cong \pi_{s_{k,m}} \otimes \mathfrak{D}_m$ occurring in $L^2(\Gamma\backslash G, \rho, (0, m))$,

$$\frac{\exp(R|1/2-s|)}{R} \ll \sum_k h(Rr_{k,m}),$$

where as usual $s_{k,m}=\frac{1}{2}+r_{k,m}$. Now use the trace formula with weight $(0,m)$ to transform this sum to a sum over conjugacy classes

$$\sum_k h(Rr_{k,m})=\frac{c_1\chi_\rho(1)(2|m|-1)}{16R^2\pi^2}\int_{\mathbb{R}} h(r)r\tanh(\pi Rr)dr$$
$$\frac{1}{R}\sum_{\{\gamma\}\in h.e} c_\gamma\chi_\rho(\gamma)\frac{g_1(\frac{l_{\gamma_1}}{R})}{\sinh(\frac{l_{\gamma_2}}{2})}\frac{ie^{i(2|m|-1)\theta_2}}{2\sin(\theta_2)}$$

In order to evaluate this sum we first add the contribution of all other wights in a window around T (thus only making it bigger). Let ψ be a positive smooth function supported on $[\frac{1}{2},\frac{5}{2}]$ with $\psi(1)=1$, and consider the sum

$$\sum_m \psi(\frac{m}{T})\sum_k h(Rr_{k,m}).$$

From the positivity of ψ this sum is still bounded from below by $\frac{\exp(R|1/2-s|)}{R}$. On the other hand if we replace the inner sum with the right hand side of the trace formula we get

$$\sum_m \psi(\frac{m}{T})\sum_k h(Rr_{k,m})=$$
$$\frac{c_1\chi_\rho(1)}{16R^2\pi^2}\int_{\mathbb{R}} h(r)r\tanh(\pi Rr)dr\sum_m(2|m|-1)\psi(\frac{m}{T})$$
$$+\frac{1}{R}\sum_{\{\gamma\}\in h.e} c_\gamma\chi_\rho(\gamma)\frac{g_1(\frac{l_{\gamma_1}}{R})}{\sinh(\frac{l_{\gamma_2}}{2})\sin(\theta_2)}\sum_m\psi(\frac{m}{T})e^{i(2m-1)\theta_2}$$

The first term is bounded by $O(\chi_\rho(1)\frac{T^2}{R^2})$ (recall we are only considering $\frac{T}{2}\leq m\leq\frac{5T}{2}$). We can bound the second term by

$$\frac{\chi_\rho(1)}{R}\sum_{\substack{\{\gamma\}\in h.e\\ l_{\gamma_1}\leq R}}|\frac{c_\gamma}{\sinh(\frac{l_{\gamma_2}}{2})\sin(\theta_2)}||\sum_m\psi(\frac{m}{T})e^{2im\theta_2}|$$

Now use Poisson summation to get

$$|\sum_m\psi(\frac{m}{T})e^{2im\theta}|=|T\sum_m\hat{\psi}(T(\theta+2m))|,$$

where $\hat{\psi}$ is the Fourier transform of ψ. From the fast decay of $\hat{\psi}$ we can deduce that the main contribution is given by $T\hat{\psi}(T\theta)$, which is bounded by $O_\epsilon(T^{-1})$ for $\theta\geq T^{-1+\epsilon}$ and by $O(T)$ for $\theta\leq T^{-1+\epsilon}$.

Thus, exactly as in the previous case we get

$$\frac{\exp(R|1/2-s|)}{R} \ll_\epsilon \dim(\rho)\Big(\frac{T^2}{R^2} + \frac{T}{R}\sum_{\substack{|t_1|\le e^{R/2} \\ 2-\frac{1}{T^{2-\epsilon}}<|t_2|<2}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}} + \frac{1}{RT}\sum_{\substack{|t_1|\le e^{R/2} \\ |t_2|\le 2}} \frac{F_\Gamma(t)}{\sqrt{(t_1^2-4)(t_2^2-4)}}\Big)$$

and setting $R = c\log T$ concludes the proof. $\square$

Theorem 4 is now reduced to the following two counting arguments:

Proposition 3.2 (First counting argument)**.** *There is a constant C (depending only on Γ) such that for any $(x_1, x_2) \in \mathbb{R}^2$*

$$\sharp\left\{t \in Tr(\Gamma) : |t_1 - x_1| \le T_1, |t_2 - x_2| \le T_2\right\} \le 1 + CT_1T_2$$

Proposition 3.3 (Second counting argument)**.**

$$\frac{F_\Gamma(t)}{\sqrt{|(t_1^2-4)(t_2^2-4)|}} \ll_\epsilon |(t_1^2-4)(t_2^2-4)|^\epsilon$$

Proof of Theorem 4. We give the proof of Theorem 4 from the two counting arguments. Assume that $\pi \cong \pi_1 \otimes \pi_2$ occur in $L^2(\Gamma\backslash G, \rho)$ and satisfies the hypothesis of Theorem 4. For fixed $c > 0$ and any $\epsilon > 0$ let ϵ_1, ϵ_2 such that $\epsilon = \epsilon_1 + c\epsilon_2$, then by Proposition 3.1 we have

$$T^{c|1/2-s_1|} \ll_\epsilon \dim(\rho)\Big(\frac{T^2}{\log(T)} + T\sum_{\substack{|t_1|\le T^{c/2} \\ |t_2|=2+O(T^{-2+\epsilon_1})}} \frac{F_\Gamma(t)}{\sqrt{|(t_1^2-4)(t_2^2-4)|}} + \frac{1}{T}\sum_{\substack{|t_1|\le T^{c/2} \\ |t_2|\le 2}} \frac{F_\Gamma(t)}{\sqrt{|(t_1^2-4)(t_2^2-4)|}}\Big).$$

The second counting argument (Proposition 3.3) together with the bound $|(t_1^2-4)(t_2^2-4)| \ll T^c$ (which holds for all pairs (t_1, t_2) appearing in the sum) gives

$$\begin{aligned} T^{c|1/2-s_1|} \ll_\epsilon \dim(\rho)\Big(&\frac{T^2}{\log(T)} \\ &+ T^{1+c\epsilon_2}\sharp\left\{t \in \mathrm{Tr}(\Gamma) : |t_1| \le T^{c/2},\ |t_2| = 2 + O(T^{-2+\epsilon_1})\right\} \\ &+ T^{-1+c\epsilon_2}\sharp\left\{t \in \mathrm{Tr}(\Gamma) : |t_1| \le T^{c/2},\ |t_2| \le 2\right\}\Big) \end{aligned}$$

Now by the first counting argument (Proposition 3.3) we get

$$T^{c|1/2-s_1|} \ll_\epsilon \dim(\rho)(\frac{T^2}{\log(T)} + T^{c/2-1+\epsilon})$$

concluding the proof. □

3.3. **Modifications for** $d>2$**.** Let $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})^d$ be a lattice (derived from quaternion algebra) and ρ a unitary representation. Assume that $\pi \cong \pi_1 \otimes \pi_2 \otimes \cdots \otimes \pi_d$ occurs in $L^2(\Gamma\backslash G, \rho)$ with $\pi_1 \cong \pi_{s_1}$ complementary series and let $J_1, J_2, J_3 \subset \{2,\ldots,n\}$ with J_1 the set of indices for which π_j is either complementary series or principal series with $r_j < 1$, J_2 the set of indices for which π_j is of principal series with $r_j > 1$ and J_3 the set of indices for which $\pi_j \cong \mathfrak{D}_{m_j}$. For $j \in J_2 \cup J_3$ let $T_j \geq 1$ be such that $r_j \in [T_j, 2T_j]$ for $j \in J_2$ and $|m_j| \in [T_j, 2T_j]$ for $j \in J_3$ and let $T = \prod_{j\in J_2\cup J_3} T_j$. With these notations the statement of Theorem 4 remains the same, that is for any $c>0$

$$T^{c|\frac{1}{2}-s_1|} \ll_\epsilon \dim(\rho)(\frac{T^2}{\log(T)} + T^{\frac{c}{2}-1+\epsilon}). \tag{3.1}$$

Theorem 2 now follows from (3.1) just as in the case of $d=2$. In order to prove the asymptotic estimate (3.1) in this setting, we apply the trace formula (and Poisson summation in the m_j variables) to the test function

$$h(r;m) = h_1(c\log(T)r_1)\prod_{j\in J_1} h_1(r_j)\prod_{j\in J_2} h_2(\frac{r_j}{T_j})\prod_{j\in J_3}\psi(\frac{m_j}{T_j}),$$

where h_1, h_2 and ψ are as in the proof of Proposition 3.1. The result then follows from the same estimates as in the proof of Proposition 3.1 (and some elementary combinatorics) together with the natural generalization of the two counting arguments above (the proofs of the counting arguments given below are for any $d \geq 2$).

4. Counting solutions

In the following section we give proofs for the two counting arguments. Let $\mathcal{A}$ be a quaternion algebra unramified in d real places, let $\mathcal{R}$ a maximal order in $\mathcal{A}$ and let $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})^d$ be the corresponding lattice.

4.1. **First counting argument.** The proof of the first counting argument is a direct result of the following estimate on the number of lattice points coming from a number field lying inside a rectangular box. Let $L/\mathbb{Q}$ be a totally real number field of degree n and $\iota_1,\ldots,\iota_n$ the different embeddings of L to $\mathbb{R}$. We then think of $\mathcal{O}_L$ as a lattice in $\mathbb{R}^n$ via the map $\mathcal{O}_L \ni t \mapsto (\iota_1(t), \ldots, \iota_n(t)) \in \mathbb{R}^n$. We show that the number of such lattice points in any box parallel to the axes is bounded by the volume of the box.

Remark 4.1. Note that if the volume of the box is large with comparison to the area of its boundary, then this result would follow from the fact that the volume of the fundamental domain of this lattice is given by the square root of the discriminant and is hence > 1. However, we are interested in particular in the case where the box is narrow in one direction and long in the other so that this type of argument will not work. Fortunately, there is a simple argument that works uniformly for all such boxes.

Lemma 4.1. *For any $B \subset \mathbb{R}^n$ a box parallel to the axes the number of lattice points in this box satisfy $|B \cap \mathcal{O}_L| \leq 1 + \mathrm{vol}(B)$.*

Proof. The only thing we will use is that for any $0 \neq t \in \mathcal{O}_L$ we have $N_{L/\mathbb{Q}}(t) \in \mathbb{Z} \setminus \{0\}$, and hence $N_{L/\mathbb{Q}}(t) > 1$. Let $T_1, \ldots, T_n > 0$ and $\vec{x} \in \mathbb{R}^n$ such that

$$B = \{t \in \mathbb{R}^n : |t_j - x_j| \leq T_j\}.$$

Now, decompose the segment $[x_1 - T_1, x_1 + T_1]$ into short segments of length $\frac{1}{cT_2 \cdots T_n}$ with $c > 2^{n-1}$. Then there are less then $2cT_1 \cdots T_n + 1$ segments (one of them might be shorter). Now, if there were more then $2cT_1 \cdots T_n + 1$ elements in $B \cap \mathcal{O}_L$, then there must be at least two elements $t \neq t'$ such that $\iota_1(t), \iota_1(t')$ lie in the same segment. Consequently, we get that $|\iota_1(t - t')| < \frac{1}{cT_2 \cdots T_n}$, and on the other hand for $j \neq 1$, $|\iota_j(t - t')| \leq 2T_j$. We thus get that $|N_{L/\mathbb{Q}}(t - t')| \leq \frac{2^{n-1}}{c} < 1$ in contradiction. We have thus shown that $|B \cap \mathcal{O}_L| < 2cT_1 \cdots T_n + 1$ for any $c > 2^{n-1}$ implying that indeed

$$|B \cap \mathcal{O}_L| \leq 2^n T_1 \cdots T_n + 1 = \mathrm{vol}(B) + 1.$$

$\square$

Proof of Proposition 3.2. Let $\Gamma \subset \mathrm{PSL}(2, \mathbb{R})^d$ be a lattice derived from a quaternion algebra over a totally real number field L. Denote by $\iota_1, \ldots \iota_n$ the different embeddings of L into $\mathbb{R}$. Let $(t_1, t_2, \ldots, t_d) = \mathrm{Tr}(\gamma) \in \mathrm{Tr}(\Gamma)$. Then there is $\alpha \in \mathcal{R}^1$ such that $\gamma_j = \iota_j(\alpha)$ for $1 \leq j \leq d$. Let $t = \mathrm{Tr}_{\mathcal{A}}(\alpha) \in \mathcal{O}_L$ then $t_j = \iota_j(t)$ for $j \leq d$ and for $j > d$ we have $\iota_j(\mathcal{R}^1) \subseteq SO(2)$ so $|\iota_j(t)| \leq 2$. Consequently, we can bound

$$\sharp \{(t_1, \ldots, t_d) \in \mathrm{Tr}(\Gamma) : \forall j \leq d, \ |t_j - x_j| \leq T_j\},$$

by the number of elements in

$$\{t \in \mathcal{O}_L : \forall j \leq d, \ |t_j - x_j| \leq T_j \text{ and } \forall j > d, \ |t_j| \leq 2\}$$

which is bounded by $1 + 2^{2n-d} T_1 T_2 \cdots T_d$. $\square$

4.2. **Arithmetic formula.** Before we proceed with the proof of the second counting argument, we give a formula for the counting function $F_\Gamma(t)$ in terms of certain arithmetic invariants (see appendix A for the related background from algebraic number theory).

Let $\alpha \in \mathcal{R}^1$ not in the center, and denote $\mathrm{Tr}_{\mathcal{A}}(\alpha) = a \in \mathcal{O}_L$ and $D = a^2 - 4$. The centralizer $\mathcal{A}_\alpha = \{\beta \in \mathcal{A} | \beta\alpha = \alpha\beta\} = L(\alpha)$ is a quadratic field extension isomorphic to $L(\sqrt{D})$ (via the map $\alpha \mapsto \frac{a+\sqrt{D}}{2}$). Let $\mathfrak{g} = \mathfrak{g}_\alpha \subset L$ be the set

$$\mathfrak{g} = \{u \in L | \exists x \in L,\ x + u\alpha \in \mathcal{R}\}.$$

Lemma 4.2. *The set $\mathfrak{g}$ is a fractional ideal containing $\mathcal{O}_L$ (i.e., $\mathfrak{g}^{-1}$ is an integral ideal). The ideal $d = d_\alpha = \mathfrak{g}_\alpha^2 D \subset \mathcal{O}_L$ is also an integral ideal.*

Proof. The first assertion is obvious. For the second part we show that any $u \in \mathfrak{g}$ satisfies $u^2 D \in \mathcal{O}_L$. Indeed, for any $u \in \mathfrak{g}$ there is $\beta = x + u\alpha \in \mathcal{R}$. Since we know that $\mathcal{N}_{\mathcal{A}}(\beta) = x^2 + u^2 + xua \in \mathcal{O}_L$ and $\mathrm{Tr}_A(\beta) = 2x + ua \in \mathcal{O}_L$, we can deduce that

$$u^2 D = (2x + ua)^2 - 4(x^2 + xua + u^2) = \mathrm{Tr}_{\mathcal{A}}(\beta)^2 - 4\mathcal{N}_A(\beta) \in \mathcal{O}_L.$$

$\square$

For D, d as above let $K = L(\sqrt{D})$ and denote by $\mathcal{O}_K$ the integers of K. Define the ring

$$\mathcal{O}_{D,d} = \left\{ \frac{t + u\sqrt{D}}{2} \in \mathcal{O}_K : d | (u^2 D) \right\}.$$

This is an order inside $\mathcal{O}_K$ [13, Proposition 5.5] and its relative discriminant over L is precisely the ideal d (see Lemma A.4).

Proposition 4.3. *Let $\alpha \in \mathcal{R}^1$, let $D = Tr_{\mathcal{A}}(\alpha)^2 - 4$ and $d = d_\alpha \subset \mathcal{O}_L$ as above. Under the map $\alpha \to \frac{1+\sqrt{D}}{2}$, the order $\mathcal{R}_\alpha = \mathcal{A}_\alpha \cap \mathcal{R}$ is mapped onto $\mathcal{O}_{D,d}$.*

Proof. Denote by $\mathcal{O}_\alpha$ the image of $\mathcal{A}_\alpha \cap \mathcal{R}$ under this map, so

$$\mathcal{O}_\alpha = \left\{ \frac{t + u\sqrt{D}}{2} \in L(\sqrt{D}) | \frac{t + u(2\alpha - a)}{2} \in \mathcal{R} \right\}.$$

The condition $\frac{t+u(2\alpha-a)}{2} \in \mathcal{R}$ implies that $t = \mathrm{Tr}_{\mathcal{A}}(\frac{t+u(2\alpha-a)}{2}) \in \mathcal{O}_L$ and that $u \in \mathfrak{g}$. Note that for any $u \in L$ we have the equivalence $d | (u^2 D) \Leftrightarrow \mathfrak{g}^2 D | (u^2)(D) \Leftrightarrow \mathfrak{g}^2 | (u)^2 \Leftrightarrow u \in \mathfrak{g}$. Hence $\mathcal{O}_\alpha \subset \mathcal{O}_{D,d}$.

For the other direction let $\frac{t+u\sqrt{D}}{2} \in \mathcal{O}_{D,d}$. In particular $u \in \mathfrak{g}$ and hence there is $\beta = x+u\alpha \in \mathcal{R}$. Let $\tilde{t} = \mathrm{Tr}_{\mathcal{A}}(\beta) \in \mathcal{O}_L$ then $\beta = \frac{\tilde{t}-ua}{2}+u\alpha$ and hence $4\mathcal{N}_A(\beta) = \tilde{t}^2 - u^2 D \in 4\mathcal{O}_L$. But from the definition of $\mathcal{O}_{D,d}$ we also know $t^2-u^2D \in 4\mathcal{O}_L$, hence $t^2-\tilde{t}^2 \in 4\mathcal{O}_L$ and $t \equiv \tilde{t} \pmod{2\mathcal{O}_L}$. Now $\frac{t+u(2\alpha-a)}{2} - \frac{\tilde{t}+u(2\alpha-a)}{2} = \frac{t-\tilde{t}}{2} \in \mathcal{O}_L \subset \mathcal{R}$, and hence $\frac{t+u(2\alpha-a)}{2} \in \mathcal{R}$ and $\frac{t+u\sqrt{D}}{2} \in \mathcal{O}_\alpha$. $\square$

Proposition 4.4. *With the above notation assume that $\iota_j(D) \in \mathbb{R}$ is positive for $j = 1, \dots, m_0$ and negative for $j = m_0+1, \dots, n$ for some $1 \le m_0 \le m$. Then $\mathcal{O}^1_{D,d}$ is a free group of rank m_0.*

Proof. See [13, proof of Theorem 5.7]. $\square$

Definition 4.5. Let $\epsilon_1, \epsilon_2, \dots, \epsilon_{m_0}$ be generators for $\mathcal{O}^1_{D,d}$. For each $j = 1, \dots, m_0$ choose one place of $L(\sqrt{D})$ above ι_j (that we also denote by ι_j). Define the regulator $\mathrm{Reg}(\mathcal{O}^1_{D,d})$ as the absolute value of the determinant of the $m_0 \times m_0$ matrix given by $a_{i,j} = \log|\iota_i(\epsilon_j)|$.

Proposition 4.6. *Let $\alpha \in \mathcal{R}^1$ and denote by $D = Tr_{\mathcal{A}}(\alpha)^2 - 4$, and $d = \mathfrak{g}^2_\alpha D$ as above. Then $\mathrm{vol}(\Gamma_\gamma \backslash G_\gamma) = \mathrm{Reg}(\mathcal{O}^1_{D,d})$ where $\gamma = \iota(\alpha)$.*

Proof. See [13, Proposition 6.1] $\square$

Proposition 4.7. *For Γ as above and $t \in \mathcal{O}_L$*

$$F_\Gamma(t) = \sum_{d|(D)} \mathrm{Reg}(\mathcal{O}^1_{D,d}) l(\mathcal{O}_{D,d}).$$

where the sum is over all ideals d such that $\frac{(D)}{d}$ is a square of an integral ideal, and $l(\mathcal{O}_{D,d})$ is the number of conjugacy classes of centralizers corresponding to $\mathcal{O}_{D,d}$.

Proof. Recall that

$$F_\Gamma(t) = \sum_{\substack{\{\gamma\}\in\Gamma^\sharp \\ \mathrm{Tr}(\gamma)=t}} \mathrm{vol}(\Gamma_\gamma \backslash G_\gamma).$$

We can assume that $t_j = \iota_j(t)$ for some $t \in \mathcal{O}_L$ and think on F_Γ as a function on $\mathcal{O}_L$. Replace the sum over conjugacy classes $\{\gamma\} \in \Gamma^\sharp$ to a sum over conjugacy classes $\{\alpha\} \in \mathcal{R}^{1\sharp}$. Next for $\gamma = \iota(\alpha)$, by proposition 4.6, we have that $\mathrm{vol}(\Gamma_\gamma \backslash G_\gamma) = \mathrm{Reg}(\mathcal{O}^1_{D,d})$ where $D = t^2 - 4$ and $d|(D)$ is the ideal corresponding to α as in proposition 4.3. Consequently we can write

$$F_\Gamma(t) = \sum_{d|(D)} \mathrm{Reg}(\mathcal{O}^1_{D,d}) \sharp \{\{\alpha\} | \mathrm{Tr}_{\mathcal{A}}(\alpha) = t,\ d_\alpha = d\}$$

where the sum is over all integral ideals $d|(D)$ such that $(D)/d$ is a square of an integral ideal. Now consider the map sending each conjugacy class $\{\alpha\}$ to the conjugacy class of its centralizer $\{\mathcal{R}^1_\alpha\}$. Note that two different elements of Γ with the same trace do not commute [13, Lemma 7.4], hence this map is a bijection of the set

$$\{\{\alpha\}|\mathrm{Tr}_{\mathcal{A}}(\alpha)=t,\ d_\alpha=d\}\,,$$

and the set of conjugacy classes of centralizers corresponding to $\mathcal{O}_{D,d}$. Consequently we have $\sharp\,\{\{\alpha\}|\mathrm{Tr}_{\mathcal{A}}(\alpha)=t,\ d_\alpha=d\}=l(\mathcal{O}_{D,d})$. $\square$

4.3. **Second counting argument.** Fix $\alpha\in\mathcal{R}^1$ (not in the center), let $K=L(\alpha)$ be the corresponding quadratic extension and let $\mathcal{O}=\mathcal{R}\cap K$. Then by Proposition 4.3 we have $\mathcal{O}\cong\mathcal{O}_{D,d}$ where $D=\mathrm{Tr}_{\mathcal{A}}(\alpha)^2-4$ and $d=d_\alpha$ as in Lemma 4.2. Note that if $\alpha'\in\mathcal{R}^1$ is conjugate (in $\mathcal{R}^1$) to α, then $D'=D$ and $d'=d$, so the corresponding rings are also the same. Recall that $l(\mathcal{O}_{D,d})$ is the number of $\mathcal{R}^1$-conjugacy classes of centralizers that correspond to $\mathcal{O}_{D,d}$. In the notation of Eichler (see [14, 46]) this is the number of $\mathcal{R}^1$-conjugacy classes of optimal embeddings of $\mathcal{O}$ into the maximal order $\mathcal{R}$. We now wish to give an upper bound for this number, or rather to the product $l(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^1_{D,d})$.

Let $\mathcal{C}(\mathcal{O})$ denote the class group (or the Picard group) of $\mathcal{O}$ and denote by $\sharp\mathcal{C}(\mathcal{O})=h(\mathcal{O})$ the class number. Let H denote the group of two sided ideals of $\mathcal{R}$ and H' denote the subgroup of all ideals generated by $\mathcal{O}$-ideals. Then $[H:H']=\prod_{\mathcal{P}|\mathfrak{d}}(1-(\frac{\mathcal{O}}{\mathcal{P}}))$ where $(\frac{\mathcal{O}}{\mathcal{P}})$ stands for Artin's symbol and $\mathfrak{d}$ denotes the discriminant of $\mathcal{A}$ over L [46, equation 47]. In particular, $[H:H']$ is bounded by a constant $c(\mathfrak{d})$ depending only on $\mathfrak{d}$.

Proposition 4.8.

$$l(\mathcal{O})\leq C_1\frac{h(\mathcal{O})}{[\mathcal{O}^*:\mathcal{O}^1\mathcal{O}_L^*]}.$$

where $C_1=c(\mathfrak{d})[\mathcal{R}^*:\mathcal{R}^1\mathcal{O}_L^*]$ is a constant depending only on the quaternion algebra.

Proof. Let κ be the number of pairs $(\mathfrak{M},\mathfrak{a})\in H/H'\times\mathcal{C}(\mathcal{O})$ such that the ideal $\mathfrak{M}\mathfrak{a}=\mathcal{R}\mu$ is principal. We then have [46, equation 45][1],

$$l(\mathcal{O})=\frac{[\mathcal{R}^*:\mathcal{R}^1\mathcal{O}_L^*]}{2[\mathcal{O}^*:\mathcal{O}^1\mathcal{O}_L^*]}\kappa.$$

Now use the bound $\kappa\leq[H:H']h(\mathcal{O})\leq c(\mathfrak{d})h(\mathcal{O})$ to conclude the proof. $\square$

[1]In [46] it is stated for K/L an imaginary extension, but the same proof holds here without changes.

Proposition 4.9.

$$l(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^1_{D,d}) \le C_2\sqrt{N_{L/\mathbb{Q}}(d)}\mathrm{Res}_{s=1}\zeta_K(s)$$

where $\zeta_K(s)$ is the Dedekind Zeta function corresponding to K and C_2 is a constant depending only on the quaternion algebra.

Proof. Denote by $\mathrm{Reg}(\mathcal{O}^*_{D,d})$ and $\mathrm{Reg}(\mathcal{O}^*_L)$ the regulators of $\mathcal{O}^*_{D,d}$ and $\mathcal{O}^*_L$ respectively. Combining the bound on $l(\mathcal{O}_{D,d})$ (Proposition 4.8), and the relation $\mathrm{Reg}(\mathcal{O}^*_{D,d}) = \frac{\mathrm{Reg}(\mathcal{O}^1_{D,d})\mathrm{Reg}(\mathcal{O}^*_L)}{[\mathcal{O}^*_{D,d}:\mathcal{O}^1_{D,d}\mathcal{O}^*_L]}$ (Proposition A.5) we get

$$l(O_{D,d})\mathrm{Reg}(O^1_{D,d}) \le C_1\frac{h(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^*_{D,d})}{\mathrm{Reg}(\mathcal{O}^*_L)}.$$

with C_1 the constant in Proposition 4.8.

Let $D_{K/L} \subseteq \mathcal{O}_L$ denote the relative discriminant of K/L, let $\mathfrak{f} = \{x \in \mathcal{O}_K | xO_K \subseteq \mathcal{O}_{D,d}\}$ denote the conductor of $\mathcal{O}_{D,d}$ and let $\mathfrak{f}_0 = \mathfrak{f} \cap \mathcal{O}_L$. We can bound (see Corollary A.1)

$$h(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^*_{D,d}) \le 2^{n+1}N_{L/\mathbb{Q}}(\mathfrak{f}_0)h(\mathcal{O}_K)\mathrm{Reg}(\mathcal{O}^*_K).$$

Now use the class number formula (see e.g., [38, Corollary 5.11])

$$h(\mathcal{O}_K)\mathrm{Reg}(\mathcal{O}^*_K) = \frac{2\sqrt{D_K}}{2^{n+m}\pi^{n-m}}\mathrm{Res}_{s=1}\zeta_K(s),$$

to get that

$$l(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^1_{D,d}) \le \frac{C_1}{\mathrm{Reg}(O^*_L)}N_{L/\mathbb{Q}}(\mathfrak{f}_0)\sqrt{D_K}\mathrm{Res}_{s=1}\zeta_K(s).$$

Finally replace $D_K = \frac{N_{L/\mathbb{Q}}(D_{K/L})}{D_L^2}$ (Proposition A.1) and $D_{K/L}f_0^2 = d$ (Proposition A.2) to conclude that

$$l(\mathcal{O}_{D,d})\mathrm{Reg}(\mathcal{O}^1_{D,d}) \le \frac{C_1}{\mathrm{Reg}(\mathcal{O}^*_L)D_L}\sqrt{N_{L/\mathbb{Q}}(d)}\mathrm{Res}_{s=1}\zeta_K(s).$$

$\square$

Proof of Proposition 3.3. By Propositions 4.7 and 4.9 we get

$$F_\Gamma(t) = \sum_{d|(D)}\mathrm{Reg}(\mathcal{O}^1_{D,d})l(\mathcal{O}_{D,d}) \ll \sum_{d|(D)}\sqrt{N_{L/\mathbb{Q}}(d)}\mathrm{Res}_{s=1}\zeta_K(s)$$

with $D = t^2 - 4$ and $K = L(\sqrt{D})$. For any $d|(D)$ we can bound $N_{L/\mathbb{Q}}(d) \le N_{L/\mathbb{Q}}(D)$ so

$$F_\Gamma(t) \ll \sqrt{N_{L/Q}(D)}\mathrm{Res}_{s=1}\zeta_K(s)\sharp\left\{a \subset \mathcal{O}_K | a^2|(D)\right\},$$

The number of ideal divisors of (D) is bounded by $O(N_{L/Q}(D)^{\epsilon})$ and for the residue of the Zeta function we have [31, Theorem 1]

$$\mathrm{Res}_{s=1}\zeta_K(s) \leq (\log(D_K))^{2n+1}.$$

Since $D_K \leq N_{L/\mathbb{Q}}(D) \ll \prod_j |(t_j^2-4)|$ indeed

$$F_\Gamma(t) \ll_\epsilon N_{L/Q}(D)^{1/2+\epsilon} \ll \prod_j |(t_j^2-4)|^{1/2+\epsilon}.$$

□

5. Application for Selbergs Zeta function

We conclude with the proof of Corollary 0.1 from Theorem 2 giving a zero free region for the Selberg Zeta functions $Z_m(s,\Gamma)$. For $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})^2$ irreducible without torsion and any $m \geq 1$ the corresponding Zeta function is given by

$$Z_m(s,\Gamma) = \prod_{\{\gamma\}^*_\Gamma} \prod_{\substack{\nu=0 \\ |i|<m}}^{\infty} (1-\epsilon_\gamma^i N(\gamma)^{-s-\nu})^{-1} \tag{5.1}$$

where the product is over all primitive conjugacy classes in Γ that are hyperbolic in the first coordinate and elliptic in the second. Using the trace formula with wight $(0,m)$ (as in section 1.4) one shows that $Z_m(s,\Gamma)$ is entire (except when $m=1$ where it has a simple pole at $s=1$) and satisfies a functional equation relating s and $1-s$. Also $Z_m(s,\Gamma)$ has trivial zeros at the negative integers and spectral zeros at the points s_k such that $\pi_{s_k}\otimes\mathfrak{D}_m$ appears in the decomposition of $L^2(\Gamma\backslash G)$. Now Theorem 2 implies that for any $t_0>5/6$ there are only finitely many $\pi_k=\pi_{s_k}\otimes\mathfrak{D}_m$ in the decomposition of $L^2(\Gamma\backslash G)$ with $1/2<s_{k,1}<t_0$. In particular for sufficiently large m_0 the half plane $\Re(s)>t_0$ is a zero free region for all the $Z_m(s,\Gamma)$'s with $m>m_0$.

Appendix A. Algebraic background

In this appendix we provide some background and collect a number of results from algebraic number theory that we have used. The main reference for this section is [38].

A.1. **Discriminants.** Let K/L be an extension of number fields and let $\mathcal{O}_K$ and $\mathcal{O}_L$ denote the corresponding rings of integers. For any basis $\{x_j\}$ of K/L the discriminant of the basis is defined as the determinant of the matrix $\mathrm{Tr}_{K/L}(x_ix_j)$. An order $\mathcal{O}\subseteq\mathcal{O}_K$ is a subring that has rank $[K:\mathbb{Q}]$ as a $\mathbb{Z}$ module. For any order $\mathcal{O}\subseteq\mathcal{O}_K$, the relative discriminant $d=d(\mathcal{O}/\mathcal{O}_L)$ is the ideal in $\mathcal{O}_L$ generated by the discriminants of all bases for K/L that lie in $\mathcal{O}$. When $\mathcal{O} = \mathcal{O}_K$ is the full ring of integers we denote $d(\mathcal{O}_K/\mathcal{O}_L) = D_{K/L}$ the relative discriminant of K/L. The relative discriminant of $K/\mathbb{Q}$ (respectively $L/\mathbb{Q}$) is a principal ideal in $\mathbb{Z}$, the generator of this ideal denoted by D_K (respectively D_L) is the discriminant of the field. We then have the following relation:

Proposition A.1.

$$D_K = D_L^{[K:L]} N_{L/\mathbb{Q}}(D_{K/L})$$

Proof. [38, Corollary 2.10] □

Assume now that K/L is a quadratic extension. Let $\mathcal{O} \subseteq \mathcal{O}_K$ be an order. The conductor of $\mathcal{O}$ is defined by

$$\mathfrak{f} = \mathfrak{f}(\mathcal{O}) = \{x \in \mathcal{O}_K | x\mathcal{O}_K \subset \mathcal{O}\}.$$

This is an ideal in $\mathcal{O}_K$ that measures how far is the order $\mathcal{O}$ from the full ring of integers. Denote by $\mathfrak{f}_0 = \mathfrak{f} \cap \mathcal{O}_L$ the ideal in $\mathcal{O}_L$ lying under it. We then have the following:

Proposition A.2. *Let $d = d(\mathcal{O}/\mathcal{O}_L)$ and $D_{K/L} = d(\mathcal{O}_K/\mathcal{O}_L)$ denote the relative discriminants of $\mathcal{O}$ and $\mathcal{O}_K$ over $\mathcal{O}_L$ respectively. Then $d = \mathfrak{f}_0^2 D_{K/L}$.*

This result is well known, however since we could not find a good reference we will include a short proof. We first need a few preparations. Fix $D \in \mathcal{O}_L$ such that $K = L(\sqrt{D})$ and define

$$\mathfrak{g} = \mathfrak{g}(\mathcal{O}, D) = \left\{ u \in L | \exists t \in \mathcal{O}_L \text{ s.t. } \frac{t + u\sqrt{D}}{2} \in \mathcal{O} \right\}.$$

Then $\mathfrak{g}$ is a fractional ideal containing $\mathcal{O}_L$ and $\mathfrak{g}^2 D$ is an integral ideal. (Note that the ideal $\mathfrak{g}$ depends on the choice of D but the product $\mathfrak{g}^2 D$ does not).

Lemma A.3. *We have $\mathcal{O} = \left\{ \frac{t+u\sqrt{D}}{2} \in \mathcal{O}_K | u \in \mathfrak{g} \right\}$.*

Proof. The inclusion $\mathcal{O} \subseteq \left\{ \frac{t+u\sqrt{D}}{2} \in \mathcal{O}_K | u \in \mathfrak{g} \right\}$ is obvious. To show the other direction assume that $\beta_1 = \frac{t_1+u\sqrt{D}}{2} \in \mathcal{O}_K$ with $u \in \mathfrak{g}$. Then there is $t_2 \in \mathcal{O}_L$ such that $\beta_2 = \frac{t_2+u\sqrt{D}}{2}$. For both $j = 1, 2$ we have that $N_{K/L}(\beta_j) = \frac{t_j^2-u^2D}{4} \in \mathcal{O}_L$. In particular $N_{K/L}(\beta_1) - N_{K/L}(\beta_2) = \frac{t_1^2-t_2^2}{4} \in \mathcal{O}_L$ so $t_1 \equiv t_2 \pmod{2\mathcal{O}_L}$. Now since the difference $\beta_1 - \beta_2 = \frac{t_1-t_2}{2} \in \mathcal{O}_L \subseteq \mathcal{O}$ and $\beta_2 \in \mathcal{O}$ then $\beta_1 \in \mathcal{O}$ as well. □

Lemma A.4. *The relative discriminant* $d = d(\mathcal{O})$ *of* $\mathcal{O}$ *over* L *satisfies* $d = \mathfrak{g}^2 D$.

Proof. By definition, the relative discriminant of $\mathcal{O}$ over L is the ideal generated by all elements of the form $\frac{(t_1u_2-t_2u_1)^2 D}{4}$ with $x_j = \frac{t_j+u_j\sqrt{D}}{2} \in \mathcal{O}$ such that (x_1, x_2) is a basis for K/L. Now notice that

$$\mathfrak{g} = \left\{ \frac{(t_1u_2 - t_2u_1)}{2} | x_j = \frac{t_j + u_j\sqrt{D}}{2} \in \mathcal{O} \right\}.$$

To see this, note that if $x_1, x_2 \in \mathcal{O}$ then the product $x_1x_2 \in \mathcal{O}$ as well implying that $\frac{(t_1u_2-t_2u_1)}{2} \in \mathfrak{g}$. For the other direction for any $u \in \mathfrak{g}$ take $x_1 = \frac{t+u\sqrt{D}}{2} \in \mathcal{O}$ and $x_2 = 1 = \frac{2+0\sqrt{D}}{2}$. Consequently, the discriminant is the ideal generated by $\{x^2D | x \in \mathfrak{g}\}$ which is precisely $\mathfrak{g}^2 D$. $\square$

Proof of Proposition A.2. Fix $D \in \mathcal{O}_L$ such that $K = L(\sqrt{D})$. Let $\mathfrak{g}_1 = \mathfrak{g}(\mathcal{O}_K, D)$ and $\mathfrak{g}_2 = \mathfrak{g}(\mathcal{O}, D)$. Then by Lemma A.4 we have that $D_{K/L} = D\mathfrak{g}_1^2$ and $d = D\mathfrak{g}_2^2$. It thus remains to show that $\mathfrak{f}_0 = \mathfrak{g}_1^{-1}\mathfrak{g}_2$. Now, by definition $\mathfrak{f}_0 = \{x \in \mathcal{O}_L | \forall \beta \in \mathcal{O}_K,\ x\beta \in \mathcal{O}\}$, and by Lemma A.3 this is the same as $\{x \in \mathcal{O}_L | \forall u \in \mathfrak{g}_1,\ xu \in \mathfrak{g}_2\} = \mathfrak{g}_1^{-1}\mathfrak{g}_2$. $\square$

A.2. **Regulators.** Let K be number field with r_1 real places and r_2 (conjugate pairs of) complex places. For any place ν of K define the corresponding norm by $\|x\|_\nu = |\nu(x)|$ when ν is real and $\|x\|_\nu = |\nu(x)|^2$ when ν is complex. Let $\mathcal{O}_K$ denote the ring of integers in K. By Dirichlet unit theorem the group of units $\mathcal{O}_K^*$ is a free group of rank $r = r_1 + r_2 - 1$. The regulator $\mathrm{Reg}(\mathcal{O}_K^*)$ of this group is the absolute value of the determinant of the matrix $a_{i,j} = (\log \|\epsilon_i\|_{\nu_j})$ where ν_j goes over r out of the $r+1$ places and $\epsilon_1, \ldots, \epsilon_r$ are generators for the group of units (this is independent of the choice of generators or places). There is a geometric interpretation of the regulator. Consider the logarithmic map from $\mathcal{O}_K^*$ to $\mathbb{R}^{d+1}$ sending

$$\epsilon \mapsto (\log(\|\epsilon\|_{\nu_1}), \ldots, \log(\|\epsilon\|_{\nu_{r+1}})).$$

Then the image of $\mathcal{O}_K$ is a lattice of rank r inside $\mathbb{R}^{r+1}$ with co-volume given by $\sqrt{r+1}\mathrm{Reg}\mathcal{O}_K^*$. If $U \subset \mathcal{O}_K^*$ is a subgroup of finite index, then it is also of the same rank and we can define the regulator of U in the same way (by taking $\epsilon_1, \ldots, \epsilon_r$ to be generators for U). We then have the relation $\mathrm{Reg}(U) = [O_K^* : U]\mathrm{Reg}(\mathcal{O}_K^*)$ (one can see this by comparing the co-volumes of the corresponding lattices).

Let K/L be a quadratic extension of a totally real number field L. Let $\mathcal{O}$ be an order in $\mathcal{O}_K$, then $\mathcal{O}^*$ is a subgroup of $\mathcal{O}_K^*$ of finite index. Denote by $\mathcal{O}^1 = \left\{x \in \mathcal{O} | N_{K/L}(x) = 1\right\}$ the group of (relative) norm one elements in $\mathcal{O}^*$. Let $[L:\mathbb{Q}]=n$ and let $1\le m\le n$ such that K has $2m$ real places and $(n-m)$ pairs of complex places. Then $\mathcal{O}_K^1$ is a free group of rank m. Let $\epsilon_1,\dots,\epsilon_m$ be generators for $\mathcal{O}^1$ and consider the $m\times m$ matrix given by $\log(\nu_i(\epsilon_j))$ where ν_i goes over the real places of K, where from every pair lying above the same place of L we take only one place. The regulator $\mathrm{Reg}(\mathcal{O}^1)$ is defined as the absolute value of the determinant of this matrix.

Note that $\mathcal{O}^1\mathcal{O}_L^*$ is a subgroup of $\mathcal{O}^*$ of finite index. We have the following relation

Proposition A.5.

$$\mathrm{Reg}(\mathcal{O}^*)=\frac{\mathrm{Reg}(\mathcal{O}^1\mathcal{O}_L^*)}{[\mathcal{O}^*:\mathcal{O}^1\mathcal{O}_L^*]}=\frac{\mathrm{Reg}(\mathcal{O}^1)\mathrm{Reg}(\mathcal{O}_L^*)}{[\mathcal{O}^*:\mathcal{O}^1\mathcal{O}_L^*]}.$$

Proof. See [10, proof of Theorem 1] □

A.3. **Class numbers.** For a number field K the Class group, $\mathcal{C}(\mathcal{O}_K)$, is the quotient of the group of all fractional ideals of $\mathcal{O}_K$ with the subgroup of principal ideals. This is a finite group and its order $h(\mathcal{O}_K)$ is the class number of K. The Class number formula relates the class number with other algebraic invariants of the number field.

Proposition A.6.

$$h(\mathcal{O}_K)=\frac{w\sqrt{D_K}}{2^{r_1+r_2}\pi^{r_2}\mathrm{Reg}(\mathcal{O}_K)}\mathrm{Res}_{s=1}\zeta_K(s),$$

where w is the number of roots of unity contained in K, D_K is the absolute discriminant, $\zeta_K(s)$ is the Dedekind zeta function, the numbers r_1,r_2 are the number of real and complex embeddings of K, and $\mathrm{Reg}(\mathcal{O}_K^*)$ is the regulator of $\mathcal{O}_K$.

Proof. See e.g., [38, Corollary 5.11] □

For an order $\mathcal{O}\subseteq\mathcal{O}_K$, the fractional ideals do not necessarily form a group (since not all ideals are invertible). However one can consider the group of all invertible ideals in $\mathcal{O}$. The Picard group of $\mathcal{O}$ is then the quotient of the group of all invertible fractional ideals of $\mathcal{O}$ with the subgroup of principal ideals. This group is also finite and its order $h(\mathcal{O})$ is called the class number of $\mathcal{O}$.

The class numbers $h(\mathcal{O})$ and $h(\mathcal{O}_K)$ are related by the following formula [38, Theorem 12.12]

$$h(\mathcal{O})=h(\mathcal{O}_K)\frac{[\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^*:(\mathcal{O}/\mathfrak{f}\mathcal{O})^*]}{[\mathcal{O}_K^*:\mathcal{O}^*]}. \tag{A.1}$$

where $\mathfrak{f} \subset \mathcal{O}_K$ is the conductor of $\mathcal{O}$. If we consider the product of the class number and the regulator we get

Proposition A.7.

$$h(\mathcal{O})\mathrm{Reg}(\mathcal{O}^*) = [\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^* : (\mathcal{O}/\mathfrak{f}\mathcal{O})^*]h(\mathcal{O}_K)\mathrm{Reg}(\mathcal{O}_K^*).$$

Proof. Use the above formula together with the relation $\mathrm{Reg}(\mathcal{O}^*) = [\mathcal{O}_K^* : \mathcal{O}^*]\mathrm{Reg}(\mathcal{O}_K^*)$. □

In the previous setup with K/L a quadratic extension and $\mathfrak{f}_0 = \mathfrak{f} \cap \mathcal{O}_L$ this leads to the following bound:

Corollary A.1.

$$h(\mathcal{O})\mathrm{Reg}(\mathcal{O}^*) \leq 2^{n+1} N_{L/\mathbb{Q}}(\mathfrak{f}_0) h(\mathcal{O}_K)\mathrm{Reg}(\mathcal{O}_K^*)$$

Proof. We need to give a bound for $[(\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^* : (\mathcal{O}/\mathfrak{f}\mathcal{O})^*]$. Consider the norm map from $\mathcal{N}_{K/L} : (\mathcal{O}/\mathfrak{f}\mathcal{O})^* \to (\mathcal{O}_L/\mathfrak{f}_0)^*$. The image of this map contains all the squares in $(\mathcal{O}_L/\mathfrak{f}_0)^*$ which is a subgroup of index bounded by 2^{n+1}. We can thus bound from below $\sharp(\mathcal{O}/\mathfrak{f}\mathcal{O})^* \geq 2^{-n-1}\sharp(\mathcal{O}_L/\mathfrak{f}_0)^*$. Consequently,

$$[(\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^* : (\mathcal{O}/\mathfrak{f}\mathcal{O})^*] \leq 2^{n+1}\frac{|(\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^*|}{|(\mathcal{O}_L/\mathfrak{f}\mathcal{O}_L)^*|} = 2^{n+1}\frac{N_{K/\mathbb{Q}}(\mathfrak{f})}{N_{L/\mathbb{Q}}(\mathfrak{f}_0)}.$$

Write $N_{K/\mathbb{Q}}(\mathfrak{f}) = N_{L/\mathbb{Q}}(N_{K/L}(\mathfrak{f}))$ and note that $\mathfrak{f}_0^2 \subset N_{K/L}(\mathfrak{f}) \subset \mathfrak{f}_0$ to get

$$[(\mathcal{O}_K/\mathfrak{f}\mathcal{O}_K)^* : (\mathcal{O}/\mathfrak{f}\mathcal{O})^*] \leq 2^{n+1}N_{L/\mathbb{Q}}(\mathfrak{f}_0).$$

□

References

[1] James Arthur, *A note on the automorphic Langlands group*, Canad. Math. Bull. **45** (2002), no. 4, 466–482, Dedicated to Robert V. Moody.

[2] A. F. Beardon, *The exponent of convergence of Poincaré series*, Proc. London Math. Soc. (3) **18** (1968), 461–483.

[3] Nicolas Bergeron and Laurent Clozel, *Spectre automorphe des variétés hyperboliques et applications topologiques*, Astérisque (2005), no. 303, xx+218.

[4] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, Annals of Mathematics Studies, 94. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980. xvii+388 pp.

[5] Jean Bourgain and Alex Gamburd, *On the spectral gap for finitely-generated subgroups of* SU(2), Invent. Math. **171** (2008), no. 1, 83–121.

[6] Robert Brooks, *The bottom of the spectrum of a Riemannian covering*, J. Reine Angew. Math. **357** (1985), 101–114.

[7] M. Burger and P. Sarnak, *Ramanujan duals. II*, Invent. Math. **106** (1991), no. 1, 1–11.

[8] L. Clozel, *Spectral theory of automorphic forms*, Automorphic forms and applications, IAS/Park City Math. Ser., vol. 12, Amer. Math. Soc., Providence, RI, 2007, pp. 43–93.

[9] ———, *Démonstration de la conjecture* τ, Invent. Math. **151** (2003), no. 2, 297–328.

[10] Antone Costa and Eduardo Friedman, *Ratios of regulators in totally real extensions of number fields*, J. Number Theory **37** (1991), no. 3, 288–297.

[11] M. Cowling, *Sur les coefficients des représentations unitaires des groupes de Lie simples*, Analyse harmonique sur les groupes de Lie (Sém., Nancy-Strasbourg 1976–1978), II, pp. 132–178, Lecture Notes in Math., 739, Springer, Berlin, 1979.

[12] Danijela Damjanović and Anatole Katok, *Local rigidity of restrictions of Weyl chamber flows*, C. R. Math. Acad. Sci. Paris **344** (2007), no. 8, 503–508.

[13] I. Efrat, *The Selberg trace formula for* $\mathrm{PSL}_2(\mathbb{R})^n$, Mem. Amer. Math. Soc. **65**.

[14] Martin Eichler, *Zur Zahlentheorie der Quaternionen-Algebren*, J. Reine Angew. Math. **195** (1955), 127–151 (1956).

[15] Alex Gamburd, Dmitry Jakobson, and Peter Sarnak, *Spectra of elements in the group ring of* $\mathrm{SU}(2)$, J. Eur. Math. Soc. (JEMS) **1** (1999), no. 1, 51–85.

[16] Dennis A. Hejhal, *The Selberg trace formula for* $\mathrm{PSL}(2, R)$*. Vol. I*, Springer-Verlag, Berlin, 1976, Lecture Notes in Mathematics, Vol. 548.

[17] ———, *The Selberg trace formula for* $\mathrm{PSL}(2, \mathbf{R})$*. Vol. 2*, Lecture Notes in Mathematics, vol. 1001, Springer-Verlag, Berlin, 1983.

[18] R. Howe, *Remarks on classical invariant theory*, Trans. Amer. Math. Soc., **313** (1989), no. 2 , 539–570.

[19] R. Howe and C. Moore *Asymptotic properties of unitary representations*, J. Funct. Anal. **32** (1979), no. 1, 72–96.

[20] M. N. Huxley, *Exceptional eigenvalues and congruence subgroups*, The Selberg trace formula and related topics (Brunswick, Maine, 1984), Contemp. Math., vol. 53, Amer. Math. Soc., Providence, RI, 1986, pp. 341–349.

[21] H. Jacquet and R. P. Langlands, *Automorphic forms on* $\mathrm{GL}(2)$, Springer-Verlag, Berlin, 1970, Lecture Notes in Mathematics, Vol. 114.

[22] A. Karrass and D. Solitar, *Note on a theorem of Schreier*, Proc. Amer. Math. Soc., **8** (1957), 696–697

[23] Henry H. Kim and Peter Sarnak, *Refined estimates towards the ramanujan and selberg conjectures*, Appendix to Henry H. Kim, J. Amer. Math. Soc. **16** (2003), no. 1, 139–183

[24] Henry H. Kim and Freydoon Shahidi, *Functorial products for* $\mathrm{GL}_2 \times \mathrm{GL}_3$ *and the symmetric cube for* GL_2, Ann. of Math. (2) **155** (2002), no. 3, 837–893, With an appendix by Colin J. Bushnell and Guy Henniart.

[25] D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494.

[26] Robert P. Langlands, *The trace formula and its applications: an introduction to the work of James Arthur*, Canad. Math. Bull. **44** (2001), no. 2, 160–209.

[27] Peter D. Lax and Ralph S. Phillips, *The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces*, J. Funct. Anal. **46** (1982), no. 3, 280–350.

[28] Jian-Shu Li, *The minimal decay of matrix coefficients for classical groups*, Harmonic analysis in China, Math. Appl., vol. 327, Kluwer Acad. Publ., Dordrecht, 1995, pp. 146–169.
[29] Jian-Shu Li and Chen-Bo Zhu, *On the decay of matrix coefficients for exceptional groups*, Math. Ann. **305** (1996), no.2, 249–270.
[30] H. Y. Loke and G. Savin, *Rank and matrix coefficients for simply laced groups*, J. Reine Angew. Math. **599** (2006), 201–216.
[31] Stéphane Louboutin, *Explicit upper bounds for residues of Dedekind zeta functions and values of L-functions at $s = 1$, and explicit lower bounds for relative class numbers of CM-fields*, Canad. J. Math. **53** (2001), no. 6, 1194–1222.
[32] A. Lubotzky, R. Phillips, and P. Sarnak, *Hecke operators and distributing points on S^2. II*, Comm. Pure Appl. Math. **40** (1987), no. 4, 401–420.
[33] Alexander Lubotzky, *Discrete groups, expanding graphs and invariant measures*, Progress in Mathematics, vol. 125, Birkhäuser Verlag, Basel, 1994, With an appendix by Jonathan D. Rogawski.
[34] Alexander Lubotzky and Dan Segal, *Subgroup growth*, Progress in Mathematics, vol. 212, Birkhäuser Verlag, Basel, 2003.
[35] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, Berlin, 1991.
[36] Calvin C. Moore, *Ergodicity of flows on homogeneous spaces*, Amer. J. Math. **88** (1966), 154–178.
[37] Henri Moscovici and Robert J. Stanton, *R-torsion and zeta functions for locally symmetric manifolds*, Invent. Math. **105** (1991), no. 1, 185–216.
[38] Jürgen Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999, Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.
[39] Hee Oh, *Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants*, Duke Math. J. **113** (2002), no. 1, 133–192.
[40] Peter Sarnak, *Notes on the generalized Ramanujan conjectures*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 659–685.
[41] Peter Sarnak and Xiao Xi Xue, *Bounds for multiplicities of automorphic representations*, Duke Math. J. **64** (1991), no. 1, 207–227.
[42] Klaus Schmidt, *Amenability, Kazhdan's property T, strong ergodicity and invariant means for ergodic group-actions*, Ergodic Theory Dynamical Systems **1** (1981), no. 2, 223–236.
[43] Atle Selberg, *Partial Zeta function*, M-L Inst. lecture notes (1995).
[44] ———, *On the estimation of Fourier coefficients of modular forms*, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 1–15.
[45] Ira Shavel, *On surfaces obtained from quaternion algebras over real quadratic fields*, Bull. Amer. Math. Soc. **82** (1976), no. 5, 727–730.
[46] Hideo Shimizu, *On discontinuous groups operating on the product of the upper half planes*, Ann. of Math. (2) **77** (1963), 33–71.

[47] Dennis Sullivan, *Related aspects of positivity in Riemannian geometry*, J. Differential Geom. **25** (1987), no. 3, 327–351.

[48] André Weil, *Algebras with involutions and the classical groups*, J. Indian Math. Soc. (N.S.) **24** (1960), 589–623 (1961).

Department of Mathematics, University of Chicago, 5734 S. University Avenue Chicago, Illinois 60637

E-mail address: kelmerdu@math.uchicago.edu

Department of Mathematics, Princeton University, Fine Hall,Washington Road, Princeton, NJ 08544 and Institute for Advanced Study ,Princeton NJ 08540

E-mail address: sarnak@Math.Princeton.edu