UNIVERSITY OF CALIFORNIA
UNIVERSITY OF CALIFORNIA SANTA CRUZ

QUANTITATIVE SPECTRAL GAP FOR THIN GROUPS OF HYPERBOLIC ISOMETRIES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

by

Michael Magee

June 2014

The Dissertation of Michael Magee
is approved:

_____
Professor Alexander Gamburd, Chair

_____
Professor Richard Montgomery

_____
Professor Jie Qing

_____
Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

Abstract

Quantitative spectral gap for thin groups of hyperbolic isometries

by

Michael Magee

Let Λ be a subgroup of an arithmetic lattice in $\mathrm{SO}(n+1,1)$. The quotient $\mathbb{H}^{n+1}/\Lambda$ has a natural family of congruence covers corresponding to ideals in a ring of integers. We establish a super-strong approximation result for Zariski-dense Λ with some additional regularity and thickness properties. Concretely, this asserts a quantitative spectral gap for the Laplacian operators on the congruence covers. This generalizes results of Sarnak and Xue (1991) and Gamburd (2002).

Acknowledgments

I thank my advisor Alex Gamburd for sharing his work with me, and for being a source of inspiration and support. I very much look forward to future collaboration with him.

Regarding this manuscript in particular, the construction which features in Section 3 is due to Professor McMullen. I also thank Lior Silberman for explaining his work to me. The referee's comments were very useful, in particular the pointer to the work of Bergeron and Clozel, and I would like to thank them for their careful reading. I am also grateful to Professor Sarnak for valuable feedback on an earlier draft.

I learned a lot from Martin Weissman and I owe a lot to his teaching. I've had many enjoyable conversations with Richard Montgomery throughout my time at Santa Cruz. Thanks also go to Jie Qing for some insights and encouragement. I'd also like to thank Sandra Yates for her assistance in administrative matters. I am lucky enough to have had many excellent teachers in the UK, without whom I would not have made it to this point. They include Graham Murphy, Gleb Gribakin, David Lyttle, Dennis Barden, Ruth Williams, Ross Lawther and Antony Wassermann. I'd like to thank my friend Alex Beloi. He has heard almost everything I have been thinking about and his inputs have been very important. I would not be in California if not for Phil Freeman, who I also thank.

I owe so much to my family. They have always supported my ambition and education. I know it has not been easy for them while I have been so far away from home and their understanding has been very important to me. Lastly, to Claire: without your love and support none of this would have been possible.

Chapter 1

Introduction

Let $n > 1$, F a totally real number field with a fixed infinite place, and $\mathcal{G} = \mathrm{SO}(F^{n+2}, q)$ the closed F-subgroup of GL_{n+2} which preserves a quadratic form q defined over F. We require $\mathcal{G}(\mathbb{R}) \cong \mathrm{SO}(n+1, 1)$ at the fixed place and compact at the other real places. Let $\mathcal{O}_F$ be the ring of integers of F and $\Gamma = \mathcal{G}(F) \cap \mathrm{GL}_{n+2}(\mathcal{O}_F)$. Ideals $\mathcal{I}$ in $\mathcal{O}_F$ give a level structure by defining

$$\Gamma(\mathcal{I}) = \{\gamma \in \Gamma : \gamma \equiv I \mod \mathcal{I}\} \tag{1.1}$$

the *principal congruence subgroup* at level $\mathcal{I}$. Let Λ be a subgroup of Γ which is Zariski-dense in $\mathcal{G}(\mathbb{C})$. We also assume that the traces of Λ in the adjoint representation generate the ring $\mathcal{O}_F$. Then Λ inherits a level structure by defining

$$\Lambda(\mathcal{I}) = \Lambda \cap \Gamma(\mathcal{I}). \tag{1.2}$$

The groups Γ, Λ, $\Lambda(\mathcal{I})$ act by isometries on hyperbolic space $\mathbb{H}^{n+1}$. On one hand, Λ can be *thin* in Γ as it is possibly infinite index, but on the other hand it is thick

enough (Zariski-dense with large trace field) so that the group structure does not degenerate. We further require that Λ is geometrically finite. This means that any Dirichlet fundamental domain for Λ in $\mathbb{H}^{n+1}$ is finitely faced, and implies that Λ is finitely generated[1].

Adding this geometric regularity gives us a further gauge of thickness. For a point $o \in \mathbb{H}^{n+1}$ the orbit Λo accumulates on a subset of the boundary S^n_∞ of $\mathbb{H}^{n+1}$. This is called the limit set of Λ and denoted $L(\Lambda)$. This set is Cantor-like and has an associated Hausdorff dimension $\delta(L(\Lambda))$. The motif of this paper is that if $\delta(L(\Lambda))$ is large enough, that is we have 'sufficient thickness', then there is a 'super-strong' approximation statement for the congruence quotients $\Lambda/\Lambda(\mathcal{I})$. Concretely, this asserts the existence of a spectral gap.

The congruence quotients $\Lambda/\Lambda(\mathcal{I})$ act as deck transformations on the quotient space $X(\mathcal{I}) \equiv \mathbb{H}^{n+1}/\Lambda(\mathcal{I})$ and induce a locally isometric covering

$$\pi_{\mathcal{I}} : X(\mathcal{I}) \to X \equiv \mathbb{H}^{n+1}/\Lambda. \tag{1.3}$$

We make some extra assumptions on Λ to ease exposition. By Selberg's Lemma [Se-60], which states that any finitely generated matrix group has a normal subgroup of finite index without torsion, we can pass to a finite index normal subgroup without elliptic or orientation reversing elements. In addition we may need to pass to a subgroup which is the kernel of the spinor norm at a localization of $\mathcal{O}_F$, we deal with this subtlety in Section 2.1. Any analysis will then occur on finite coverings of the initial X and $X(\mathcal{I})$.

It follows in Section 2.1 from the work of Weisfeiler [Weis-84] that away from

[1] We refer the reader to [Ka-00] for general details about hyperbolic geometry.

finitely many primes the natural inclusion $\Lambda/\Lambda(\mathcal{I}) \hookrightarrow \Gamma/\Gamma(\mathcal{I})$ is *onto* some large subgroup[2] $(\Gamma/\Gamma(\mathcal{I}))'$. In other words, if we choose a set of generators $S = \{A_1, \ldots, A_k\}$ for Λ and consider the Cayley graph $\mathcal{H}_{\mathcal{I}} \equiv \mathcal{H}(S\Gamma(\mathcal{I}), (\Gamma/\Gamma(\mathcal{I}))')$ then this graph is *connected*. This is the initial strong approximation statement which is to be strengthened.

A natural such strengthening is to insist that these graphs be 'highly connected', a concept which can be made precise by defining the expansion coefficient of a k-regular graph $\mathcal{H}$

$$c(\mathcal{H}) = \inf \left\{ \frac{|N(W)|}{|W|} \ : \ |W| < \frac{1}{2}|\mathcal{H}| \right\} \tag{1.4}$$

where W runs over subsets of the vertex set of $\mathcal{H}$ and $N(W)$ denotes its set of neighbours. The Cayley graphs in question have associated discrete Laplacians which are neighbour-averaging operators on functions on the vertices. As in [Lu-93] the spectral theory of these operators is related to the expansion coefficients of the graphs. The highly-connectedness will then be a property of a family of graphs (in our case $\mathcal{H}_{\mathcal{I}}$), we say that $\mathcal{H}_{\mathcal{I}}$ are a family of expanders if there exists C real such that

$$\liminf_{|\mathcal{O}_F/\mathcal{I}| \to \infty} c(\mathcal{H}_{\mathcal{I}}) \geq C > 0. \tag{1.5}$$

Suitably reinterpreted this asserts a spectral gap for the graph Laplacians. It is a fairly direct consequence[3] of Fell's continuity of induction [F-62] that by passing through representation theoretic descriptions of the action of the Laplacians (graph theoretic and diffeo-geometric), it is sufficient to prove a spectral gap result 'up above' to establish

[2]The kernel of a spinor norm at a finite semi-local ring.

[3]This argument appears in detail in [Ga-02].

the expansion property for the Cayley graphs $\mathcal{H}_{\mathcal{I}}$. To summarize, the spectral gap for the manifolds $X(\mathcal{I})$ would imply a connectedness property which naturally goes beyond strong approximation, hence 'super-strong' approximation.

To see the asserted spectral gap, we consider the Laplacian operators Δ_X, $\Delta_{X(\mathcal{I})}$, and in particular, their L^2 spectra denoted $\Omega(X)$, $\Omega(X(\mathcal{I}))$ respectively. The bottom of the spectrum was characterized by Sullivan in [Sul-82], and Lax and Phillips established a finiteness property in [LP-82]. These results can be summed up as follows.

1.0.1 Theorem (Sullivan, Lax and Phillips) *Suppose $\delta > n/2$. The following hold:*

1. *The bottom of the L^2 spectrum of the Laplacian on X (resp. $X(\mathcal{I})$) is an eigenvalue of multiplicity one at $\lambda_0 = \delta(n-\delta)$.*

2. *The L^2 spectrum in the range $[\delta(n-\delta), n^2/4)$ consists of finitely many discrete eigenvalues.*

The hypothesis of our main theorem will imply $\delta > n/2$.

The Hausdorff dimension of the limit set $L(X(\mathcal{I}))$ is the same for all $\mathcal{I}$, so this result tells us the bottom of the spectrum at all levels. Moreover any eigenfunction of Δ_X lifts to an eigenfunction of $\Delta_{X(\mathcal{I})}$ with the same eigenvalue, so by the finiteness statement in the previous theorem we know that at level $\mathcal{I}$ the discrete spectrum of $\Delta_{X(\mathcal{I})}$ in $(0, n^2/4)$ consists of that of Δ_X in addition to finitely many new eigenvalues. Our main theorem gives an explicit range in which there can be no new eigenvalues (for $\mathcal{I}$ avoiding finitely many primes). Together with the finiteness of the spectrum

in $(0, n^2/4)$ for X and at each of the finitely many excluded levels, this implies the existence of a spectral gap.

Before stating our main theorem we give some history of the spectral gap. Our starting point is Selberg's seminal paper [Se-65] where it is proved

1.0.2 Theorem (Selberg) *Let $\Gamma(N)$ be the principal congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$ at level N, and $\Gamma' \supset \Gamma(N)$. Then letting $X'(N) = \mathbb{H}^2/\Gamma'$ and writing $\lambda_1(X'(N))$ for the first non-zero eigenvalue of $\Delta_{X'(N)}$ we have for any $N \geq 1$*

$$\lambda_1(X'(N)) \geq 3/16. \tag{1.6}$$

It was conjectured by Selberg at the same time that in fact, with notation as before

$$\lambda_1(X'(N)) \geq 1/4. \tag{1.7}$$

Selberg's 3/16 result is no longer the state of the art[4], however, the conjectured 1/4 remains unattained. This is a fundamental open problem of modular forms. The reader can read Sarnak's notice [Sa-95] for a friendly exposition of the subject, as well as the notes of Sarnak [Sa-05] for more recent developments.

Theorem 1.0.2 was generalized to three dimensions by Sarnak [Sa-83] where it was proved that if E is any quadratic imaginary number field and $\mathcal{O}_E$ the ring of integers then

$$\lambda_1(\mathbb{H}^3/\mathrm{SL}_2(\mathcal{O}_E)) \geq 3/4. \tag{1.8}$$

[4]Luo, Rudnick and Sarnak in [LRS-95] proved $\lambda_1(X'(N)) \geq 171/784$ by using properties of GL_3 Rankin-Selberg L-functions. Shortly after this Iwaniec [I-96] proved the slightly weaker $\lambda_1(X'(N)) \geq 10/49$ by using only the GL_2 theory. Kim and Shahidi [KiSh-02] proved $\lambda_1(X'(N)) \geq 66/289$ via the existence of the functorial symmetric cube for GL_2. After some further developments along these lines (functorial powers) Kim and Sarnak [KiSa-03] proved the current best result, which says $\lambda_1(X'(N)) \geq 975/4096$.

The method used there is very much in the spirit of Selberg's proof of (1.6). It extends to congruence subgroups of $\mathrm{SL}_2(\mathcal{O}_E)$. This was further extended to arbitrary dimension by Elstrodt, Grunewald and Mennicke [EGM-90] and Cogdell, Li, Piatetski-Shapiro and Sarnak [CLPS-91] independently. Both these papers prove that if Q is a quadratic form of signature $(1, n+1)$, $n > 1$, Q isotropic over $\mathbf{Q}$, and Γ a congruence subgroup of $\mathrm{SO}^0_{n+2}(\mathbf{Z}, Q)$ then

$$\lambda_1(\mathbb{H}^{n+1}/\Gamma) \geq \frac{2n-1}{4}. \tag{1.9}$$

Work of Burger and Sarnak [BS-91] gave further progress by allowing one to link the Laplacian spectrum of conguence hyperbolic manifolds to the automorphic spectrum of GL_2. Together with results of Blomer and Brumley [BB-11] this lifting argument yields that when $\mathcal{G}$ and Γ are as in our setup and $\Gamma(\mathcal{I})$ is a congruence subgroup,

$$\lambda_1(\mathbb{H}^{n+1}/\Gamma(\mathcal{I})) > \frac{25}{32}\left(n - \frac{25}{32}\right). \tag{1.10}$$

Recently, following the proof of the fundamental lemma by Ngô [N-10] and the weighted fundamental lemma by Chaudouard and Laumon [CL-10, CL-09], certain conditional results of Arthur appearing in [A-05] have become fact. Using these results of Arthur, Bergeron and Clozel proved in [BC-12] the following result on the spectrum of the Laplacian.

1.0.3 Theorem (Bergeron, Clozel) *Let $\mathcal{G}$ be a $\mathbf{Q}$-group obtained by restriction of scalars from a special orthogonal group (split or quasi-split) over a totally real number field. Additionally suppose that $\mathcal{G}$ does not come from a twisted form ${}^3D_4$ or ${}^6D_4$ and that $\mathcal{G}(\mathbb{R})$ is the product of $\mathrm{SO}(n+1,1)$ with a compact group. For any torsion free*

congruence subgroup $\Gamma \subset \mathcal{G}$*, the spectrum of the Laplacian on* $\mathbb{H}^{n+1}/\Gamma$ *is contained in the set*

$$\bigcup_{0 \leq j < \frac{n}{2}} \{j(n-j)\} \cup \left[\frac{n^2}{4} - \left(\frac{1}{2} - \frac{1}{N^2+1} \right)^2, \infty \right), \tag{1.11}$$

where $N = n+1$ *if* n *is odd and* $N = n+2$ *if* n *is even. In particular, when* $n \geq 3$ *we have the spectral gap result*

$$\lambda_1(\mathbb{H}^{n+1}/\Gamma) \geq n-1. \tag{1.12}$$

This result feeds our main theorem via the work of Kelmer and Silberman [KeSi-10] relating the spectral theory to the lattice point count. When $n \geq 3$ Theorem 1.0.3 gives the best possible input for our method. When $n = 2$, we use the lattice point of Sarnak and Xue [SX-91] which is also the best possible.

The proof of Theorem 1.0.2, for example, relies essentially on the underlying arithmetic of the modular group $\mathrm{SL}_2(\mathbf{Z})$ and associated Kloosterman sums[5]. We will not have access to such rich arithmetic and will rely on a more robust 'almost geometric' method developed by Sarnak and Xue [SX-91] and extended to the two dimensional infinite volume case by Gamburd [Ga-02]. Sarnak and Xue proved

1.0.4 Theorem (Sarnak, Xue) *Let* F *be a totally real number field with a fixed infinite place, and* $\mathcal{O}_F$ *the ring of integers of* F*. Let* $\mathcal{G}$ *be an orthogonal* F*-subgroup of* GL_4 *with* $\mathcal{G}(\mathbb{R}) \cong \mathrm{SL}_2(\mathbb{R})$ *(resp.* $\mathrm{SL}_2(\mathbb{C})$*) at the fixed place and compact at the other real places. Let* Γ *be a finite index subgroup of* $\mathcal{G}(F) \cap \mathrm{GL}_4(\mathcal{O}_F)$ *which is cocompact*

[5]The estimates for Kloosterman sums come from Weil [Weil-48] and appeal to the Riemann hypothesis for curves (also proved by Weil). Iwaniec proved (1.6) in [I-89] using Kloosterman sums but without relying on Weil's bound. Gelbart and Jacquet [GJ-78] proved that 3/16 is not attained in (1.6) by a very different method to that of Selberg.

in $\mathrm{SL}_2(\mathbb{R})$ *(resp.* $\mathrm{SL}_2(\mathbb{C})$*). Then for large enough prime ideals* $\mathcal{P} \subset \mathcal{O}_F$

$$\Omega(\mathbb{H}^{n+1}/\Gamma(\mathcal{P})) \cap [0, \mu) = \Omega(\mathbb{H}^{n+1}/\Gamma(1)) \cap [0, \mu)$$

where $n = 1$ *(resp.* 2*) and* $\mu = 5/36$ *(resp.* $11/36$*) in the case of* $\mathrm{SL}_2(\mathbb{R})$ *(resp.* $\mathrm{SL}_2(\mathbb{C})$*).*

The Sarnak and Xue machine makes use of the fact that if new eigenvalues appear, they are of high multiplicity. This follows by bounding below the dimension of new irreducible representations of the factor group $\Lambda/\Lambda(\mathcal{I})$. Everything we need in this direction is contained in Section 2.1. The multiplicities feature in one side of the trace formula, and the other side of the trace formula can be related to a lattice point count by choosing the right family of automorphic kernels to trace. We introduce the necessary kernels and gather some estimates on the lattice point count and spherical functions in Section 2.3.

In the cocompact case the lattice point count plays against the multiplicity estimate via the trace formula to give a contradiction when new eigenvalues appear in a certain range. This is the approach of Sarnak and Xue. However, in the infinite volume case the trace formula does not hold as is, and must be reinterpreted as an inequality. Further repairs are needed and these were made by Gamburd in the two dimensional ($\mathbb{H}^2$) case to prove in [Ga-02]

1.0.5 Theorem (Gamburd) *Let* $\Lambda = \langle A_1, \ldots, A_k \rangle$ *be a finitely generated subgroup of* $\mathrm{SL}_2(\mathbf{Z})$ *with* $\delta > 5/6$*. Let* $X(p) = \mathbb{H}^2/\Lambda(p)$*. For* p *large enough*

$$\Omega(X(p)) \cap [\delta(1-\delta), 5/36) = \Omega(X(1)) \cap [\delta(1-\delta), 5/36).$$

The 'Collar Lemmas' in [Ga-02] form a key part of the generalization of Sarnak and Xue's method to infinite volume. Roughly speaking, these state that eigenfunctions corresponding to eigenvalues $< 1/4$ in $X(p)$ are uniformly bounded through p away from concentrating near infinity. The methods used to prove these do not obviously generalize to higher dimensions. The needed generalization is the thrust of this paper and appears in Section 2.2 as Lemma 2.2.8 along with the prerequisite geometry.

The proof of Lemma 2.2.8 has a nice heuristic as follows. If one considers classical motion of a particle on a line under a step potential of height V_0, and the conserved energy E is $< V_0$, then the particle will never enter the region covered by the step. This is due to $E = K + V$ and $K \geq 0$. In the quantum mechanical version of the same system it is no longer true that the stationary wave function is zero inside the step (quantum tunnelling). However provided $E < V_0$ is bounded away from V_0 we should get uniform bounds through E which say the wave function cannot be arbitrarily concentrated inside the step. The positive Laplacian $\Delta_{X(\mathcal{I})}$ plays the role of a Schrödinger operator for free dynamics on $X(\mathcal{I})$, the eigenvalues of $\Delta_{X(\mathcal{I})}$ corresponding to energy levels. We seek uniform bounds on eigenfunctions with eigenvalues bounded away from and less than $n^2/4$, which can be thought of as the escape energy, so that these eigenfunctions are bound states. Given that their energy is uniformly bounded through $\mathcal{I}$ away from escape, they should not concentrate near infinity. Making formal sense of this argument constitutes the bulk of the proof.

All the machinery is brought together in Section 2.4 to prove the following.

1.0.6 Theorem (Main Theorem) *Let F be a totally real number field with a fixed infinite place, and $\mathcal{O}_F$ the ring of integers of F. Let $\mathcal{G} = \mathrm{SO}(F^{n+2}, q)$ the closed F-subgroup of GL_{n+2} which preserves a quadratic form q defined over F. Assume $\mathcal{G}(\mathbb{R}) \cong \mathrm{SO}(n+1,1)$ at the fixed place and compact at the other real places. Let Λ be a subgroup of $\mathcal{G}(F) \cap \mathrm{GL}_{n+2}(\mathcal{O}_F)$ with the following properties.*

Algebraic Fullness *Λ is Zariski-dense in $\mathcal{G}(\mathbb{C})$ and the traces of Λ generate $\mathcal{O}_F$.*

Geometric Regularity *The image of Λ at the fixed place is geometrically finite, orientation preserving and torsion free as an isometry group of $\mathbb{H}^{n+1}$.*

Fractal Fullness *The Hausdorff dimension δ of the limit set of Λ is greater than s_n^0, defined*

$$s_n^0 \equiv n - \frac{2(n-1)}{(n+1)(n+2)}. \tag{1.13}$$

Let $X(\mathcal{I}) = \mathbb{H}^{n+1}/\Lambda(\mathcal{I})$ for $\mathcal{I}$ an ideal in $\mathcal{O}_F$. Then by replacing Λ with a finite index subgroup if necessary we have for $|\mathcal{O}_F/\mathcal{I}|$ large enough and $\mathcal{I}$ coprime to a finite set of primes

$$\Omega(X(\mathcal{I})) \cap \left[\delta(n-\delta), s_n^0(n-s_n^0)\right) = \Omega(X(1)) \cap \left[\delta(n-\delta), s_n^0(n-s_n^0)\right). \tag{1.14}$$

The finite index subgroup of Λ which we may need to pass to is the Λ_1 of Section 2.1. This is not necessary if one assumes that Λ is contained in the kernel of the appropriate spinor norm. The index of this subgroup has independent bounds which depend on the number of generators of Λ and the structure of a localization[6] $(\mathcal{O}_F)_S$ respectively.

[6]This localization is given by strong approximation.

1.0.7 Corollary (Main Corollary) *For Λ as before (replace Λ with the finite index spinor norm kernel if necessary) and $\delta > s_n^0$, $\Omega(X(\mathcal{I}))$ has a spectral gap. That is to say, writing $\lambda_1(X(\mathcal{I}))$ for the second smallest eigenvalue of $\Delta_{X(\mathcal{I})}$, for $|\mathcal{O}_F/\mathcal{I}|$ large and $\mathcal{I}$ coprime to a finite set of primes*

$$\lambda_1(X(\mathcal{I})) \geq \min\left(\lambda_1(X), s_n^0(n - s_n^0)\right).$$

Some remarks are due before we mention applications. When $\Lambda = \Gamma$ we obtain the result of Clozel [Clo-03] on property (τ), albeit with a weaker bound. For any arithmetic lattice the Borel Density Theorem [Bor-60] implies Zariski-density, so our result also applies (when our geometric criteria are met). In this case we obtain the current best bound for an arbitrary arithmetic lattice in $\mathrm{SO}(n+1,1)$. This is due to a somewhat trivial tightening in Lemma 2.1.3 of the arguments in Kelmer and Silberman [KeSi-10] together with a nontrivial improvement in Lemma 2.1.4. When Λ is of infinite index our result is entirely new. We show that our result is not vacuous in this case by constructing eligible Λ in Section 3.

One important application of the spectral gap is the Bourgain-Gamburd-Sarnak affine linear sieve introduced in [BGS-06], [BGS-10], which we recall now.

1.0.8 Theorem (Bourgain, Gamburd, Sarnak) *Let $G \subset \mathrm{GL_n}$ be a connected, simply connected, absolutely simple algebraic group defined over $\mathbf{Q}$. Let $f \in \mathbf{Q}[G]$ be a non-zero non-unit with t irreducible factors in $\mathbf{Q}[G]$. Let Λ be a subgroup of $G(\mathbf{Q}) \cap \mathrm{GL}_n(\mathbf{Z})$ finitely generated by a set S. We suppose the pair (Λ, f) has the following properties.*

Algebraic Fullness *Λ is Zariski-dense in G.*

No local congruence obstructions *For all integer $q \geq 2$ there exists $x \in \Lambda$ with*

$$(f(x), q) = 1.$$

Square free Expansion *As q runs through square free integers the Cayley graphs*

$$\mathcal{H}(S\Lambda(q), \Lambda/\Lambda(q)) \text{ form an expander family.}$$

Then there exists r such that the set of $x \in \Lambda$ such that $f(x)$ has at most r prime factors is Zariski-dense in G. Moreover the minimal such r is bounded explicitly and effectively in terms of the spectral gap in the expander family.

This result is given in [BGS-10, Theorem 1.6]. Using tools of additive combinatorics Bourgain and Gamburd [BG-08] established the expansion property for Zariski-dense $\Lambda \subset \mathrm{SL}_2(\mathbf{Z})$ through prime levels. In [BGS-09] the expansion property in $\mathrm{SL}_2(\mathbf{Z})$ for Zariski-dense Λ is proved for square free levels, and an equivalence between expansion in Cayley graphs and the spectral gap for the spaces $\mathbb{H}^2/\Lambda(q)$ is given. In case $\delta(\Lambda)$ is $\leq 1/2$ there is no discrete L^2 spectrum and the gap has to be interpreted as a pole free region of the meromorphically continued resolvent $(\Delta_{\mathbb{H}^2/\Lambda(q)} - s(1-s))^{-1}$. Affine sieve methods are used to sharply estimate the quantity

$$|\{x \in \Lambda \,:\, |x| \leq T, \text{ all the irreducible factors of } f \text{ have prime evaluation at } x\}|$$

using the non-explicit gap. Furthermore (still in [BGS-09]) it is shown that there is an r such that

$$|\{x \in \Lambda \,:\, |x| \leq T,\ f(x) \text{ has at most } r \text{ prime factors}\}| \tag{1.15}$$

has a good bound below, in particular implying that $f(\Lambda)$ contains infinitely many r-almost primes, that is, numbers which are products of at most r primes. This r can

be determined explicitly using either the value of the L^2 spectral gap or the size of the pole free region of the (continued) resolvent (corresponding to $\delta > 1/2$ and $\delta \leq 1/2$ respectively).

The explicit gap is utilized in the paper of Kontorovich [Ko-09]. There the affine sieve theory is applied (with the necessary adaptations) to the function

$$f(c,d) = c^2 + d^2$$

and the orbit $\mathcal{O} = (0,1)\Gamma$, for Γ an infinite index, Zariski-dense, finitely generated subgroup of $\mathrm{SL}_2(\mathbf{Z})$ with $\delta(\Gamma) > 149/150$ and containing parabolics. Then using Gamburd's explicit 5/6 gap from [Ga-02], it is proved that $f(\mathcal{O})$ contains infinitely many 25-almost primes. Similar methods are applied (in particular also using Gamburd's 5/6 gap) by Kontorovich and Oh [KoOh-10] to the Pythagorean orbit $\mathcal{O} = (3,4,5)\Gamma$ for Γ a finitely generated Zariski-dense subgroup of $\mathrm{SO}_Q(\mathbf{Z})$,

$$Q(\mathbf{x}) = x^2 + y^2 - z^2.$$

They consider hypotenuse ($F(\mathbf{x}) = z$), area ($F(\mathbf{x}) = xy/12$) and product ($F(\mathbf{x}) = xyz/60$) functions. The affine linear sieve gives infinitely many R-almost primes in $F(\mathcal{O})$ for explicit R provided $\delta(\Gamma)$ is large and there are no local congruence obstructions for the pair $(\mathcal{O}, F)$ (R and δ depend on the function F considered).

Theorem 1.0.6 will yield similar applications via the affine linear sieve.

1.0.1 Notation

Throughout this thesis we use $f \ll g$ to mean that $f \leq Cg$ for some constant C and $f \approx g$ to mean that $g \ll f \ll g$. If subscripts are present, e.g. $\ll_\epsilon$, this indicates

that the implied constant depends on the subscripts. We view the number field F and the groups $\mathcal{G}$ and Λ as fixed throughout, so all our implied constants possibly depend on these objects.

Chapter 2

Proof of the Theorem

2.1 Algebra

2.1.1 Strong Approximation and reduction to prime powers

The aim of this section is to control the factor groups $\Lambda/\Lambda(\mathcal{I})$ and their representation theory. The group $\Lambda/\Lambda(\mathcal{I})$ is naturally a subgroup of $\Gamma/\Gamma(\mathcal{I})$. In [Ga-02] the case $n = 1$ is treated at prime levels. In this case enough is known about the maximal subgroups of $\mathrm{SL}_2(\mathbb{F}_p)$ to provide an ad hoc proof that outside of finitely many primes p, $\Lambda/\Lambda(p) \cong \mathrm{SL}_2(\mathbb{F}_p)$. The needed bound on the nontrivial representations is that of Frobenius.

For us the description of the factor groups will follow from the work of Weisfeiler [Weis-84]. This describes strong approximation for Zariski-dense subgroups of algebraic groups which amongst other things are *simply connected*. We therefore need to carefully deal with the fact that SO is covered by Spin. The bounds on representa-

tions at prime levels are due to Seitz and Zalesskii [SZ-93] and at the general level we use a result of Kelmer and Silberman [KeSi-10] together with an improvement of our own.

Recall that F is a totally real number field with a fixed infinite place, and $\mathcal{G} = \mathrm{SO}(F^{n+2}, q)$ is the closed F-subgroup of GL_{n+2} which preserves a quadratic form q defined over F. We assume $\mathcal{G}(\mathbb{R}) \cong \mathrm{SO}(n+1, 1)$ at the fixed place and is compact at the other real places. The ring of integers of F is $\mathcal{O}_F$ and $\Gamma = \mathcal{G}(F) \cap \mathrm{GL}_{n+2}(\mathcal{O}_F)$. We drop the dependence on F and simply write $\mathcal{O} = \mathcal{O}_F$ in this section. Then we take Λ a subgroup of Γ which is Zariski-dense in $\mathcal{G}(\mathbb{C})$ and such that the traces of Λ in the adjoint representation generate $\mathcal{O}_F$. The congruence subgroups $\Gamma(\mathcal{I})$, $\Lambda(\mathcal{I})$ are kernels of the reduction map mod $\mathcal{I}$ an ideal in $\mathcal{O}$.

Compactness of $\mathcal{G}(\mathbb{R}_v)$ away from the fixed place implies discreteness of Γ at the fixed place, by using the isomorphism $\mathcal{G}(\mathbb{R}) \cong \mathrm{SO}(n+1, 1)$ we therefore realize Γ and Λ as discrete isometry groups of $n+1$ dimensional hyperbolic space $\mathbb{H}^{n+1}$. We assume that Λ is geometrically finite, hence finitely generated.

As in Weisfeiler [Weis-84, Theorem 1.1] there exists a finite set of primes S such that $\mathcal{G}$ can be given the structure of a group scheme over the localization $\mathcal{O}_S$ of $\mathcal{O}$ away from S and Λ is contained in $\mathcal{G}_{\mathcal{O}_S}(\mathcal{O}_S)$. This results from 'clearing denominators' in the definition of q, we take S as in [Weis-84, Theorem 1.1]. We have then an 'orthogonal $\mathcal{O}_S$-module' in the sense of Bass [Bass-74] by equipping $P = \mathcal{O}_S^{n+2}$ with $q : \mathcal{O}_S^{n+2} \to \mathcal{O}_S$, and $\mathcal{G} = \mathrm{SO}(P, q)$ as a group scheme over $\mathcal{O}_S$. Following Bass

[Bass-74] there is a short exact sequence of group schemes (suppressing q)

$$1 \to \mu_2 \to \mathrm{Spin} \to \mathcal{G}_{\mathcal{O}_S} \to 1 \tag{2.1}$$

which is exact in the *fppf*[1] topology on $\mathrm{Spec}(\mathcal{O}_S)$.

This yields the sequence in cohomology

$$1 \to \mu_2(\mathcal{O}_S) \to \mathrm{Spin}(\mathcal{O}_S) \to \mathcal{G}_{\mathcal{O}_S}(\mathcal{O}_S) \to^{\sigma} \mathrm{H}^1(\mathrm{Spec}(\mathcal{O}_S), \mu_2). \tag{2.2}$$

There is an isomorphism $H^1(\mathrm{Spec}(\mathcal{O}_S), \mu_2) \cong \mathrm{Discr}(\mathcal{O}_S)$ which converts σ into the spinor norm SN. The discriminant group $\mathrm{Discr}(\mathcal{O}_S)$ fits into the exact sequence

$$0 \to \mu_2(\mathcal{O}_S) \to \mathcal{O}_S^* \to^2 \mathcal{O}_S^* \to \mathrm{Discr}(\mathcal{O}_S) \to \mathrm{Pic}(\mathcal{O}_S) \to^2 \mathrm{Pic}(\mathcal{O}_S), \tag{2.3}$$

which implies that $\mathrm{Discr}(\mathcal{O}_S)$ is a finite abelian group of exponent 2.

We let $\tilde{\Lambda}$ denote the preimage of Λ in $\mathrm{Spin}(\mathcal{O}_S)$. The Strong Approximation theorem of Weisfeiler [Weis-84] then states that there is a finite index subgroup $\tilde{\Lambda}_0$ of $\tilde{\Lambda}$ such that the image of $\tilde{\Lambda}_0$ is dense in the group $\mathrm{Spin}(\hat{\mathcal{O}_S})$, where $\hat{\mathcal{O}_S}$ is the profinite completion of $\mathcal{O}_S$. In particular for $\mathcal{I}$ avoiding S the reduction map

$$\tilde{\Lambda}_0 \to \mathrm{Spin}(\mathcal{O}_S/\mathcal{I}) \cong \mathrm{Spin}(\mathcal{O}/\mathcal{I})$$

is onto.

By appealing to commutativity of the diagram obtained by reducing the sequence (2.2) modulo $\mathcal{I}$ we get that

$$\Lambda/\Lambda(\mathcal{I}) \supseteq \mathrm{Image}(\phi_{\mathcal{I}}) = \ker\left(SN_{\mathcal{I}} : \mathcal{G}_{O_S}(\mathcal{O}/\mathcal{I}) \to \mathrm{Disc}(\mathcal{O}/\mathcal{I})\right), \tag{2.4}$$

[1] faithfully flat and finitely presented

where $\phi_{\mathcal{I}} : \mathrm{Spin}(\mathcal{O}/\mathcal{I}) \to \mathcal{G}_{\mathcal{O}_S}(\mathcal{O}/\mathcal{I})$ is the covering map of finite groups and $SN_{\mathcal{I}}$ is the spinor norm at $\mathcal{I}$. If $\Lambda/\Lambda(\mathcal{I})$ is bigger than the image then there will be nontrivial representations of $\Lambda/\Lambda(\mathcal{I})$ which factor through the quotient $(\Lambda/\Lambda(\mathcal{I}))/\mathrm{Image}(\phi_{\mathcal{I}})$. We let

$$\Lambda_1 = \ker SN|_{\Lambda} \tag{2.5}$$

be the kernel of the spinor norm restricted to Λ, this is a finite index normal subgroup. The index $[\Lambda : \Lambda_1]$ is bounded independently by the size of $\mathrm{Discr}(\mathcal{O}_S)$ and by 2^L where L is the number of generators of Λ. Then we have the precise strong approximation statement

$$\Lambda_1/\Lambda_1(\mathcal{I}) = \ker(SN_{\mathcal{I}}) \tag{2.6}$$

for $\mathcal{I}$ avoiding S.

If $\mathcal{I}$ has prime factorization

$$\mathcal{I} = \prod_{i=1}^{l} \mathcal{P}_i^{r_i}$$

then the group $\Lambda_1/\Lambda_1(\mathcal{I})$ splits as a product

$$\Lambda_1/\Lambda_1(\mathcal{I}) \cong \prod_{i=1}^{l} \Lambda_1/\Lambda_1(\mathcal{P}_i^{r_i}), \tag{2.7}$$

so that bounds on the size of $\Lambda_1/\Lambda_1(\mathcal{I})$ will follow from bounds at prime power levels via $|\Lambda_1/\Lambda_1(\mathcal{I})| = \prod_i |\Lambda_1/\Lambda_1(\mathcal{P}_i^{r_i})|$. Let $\rho : \Lambda_1/\Lambda_1(\mathcal{I}) \to \mathrm{Aut}(V)$ be a nontrivial irreducible representation of level $\mathcal{I}$, i.e. ρ does not factor through a representation of $\Lambda_1/\Lambda_1(\mathcal{I}')$ for any $\mathcal{I}'|\mathcal{I}$, $\mathcal{I}' \neq \mathcal{I}$. Then ρ is a tensor product of irreducible representations $\rho_i : \Lambda_1/\Lambda_1(\mathcal{P}_i^{r_i}) \to \mathrm{Aut}(V_i)$ which are of level $\mathcal{P}_i^{r_i}$ respectively and

$$\dim \rho = \prod_i \dim \rho_i. \tag{2.8}$$

We have now reduced the needed argument to prime power level. We deal with the prime case first.

2.1.2 Prime case

Writing $k_{\mathcal{P}} = \mathcal{O}_F/\mathcal{P}$ for the residue field at $\mathcal{P}$, the previous discussion says that for $\mathcal{P}$ avoiding S we have

$$\Lambda_1/\Lambda_1(\mathcal{P}) = \begin{cases} \Omega_{\pm}(2m, |k_{\mathcal{P}}|) & \text{if } n = 2m-2 \text{ is even} \\ \\ \Omega(2m+1, |k_{\mathcal{P}}|) & \text{if } n = 2m-1 \text{ is odd.} \end{cases} \tag{2.9}$$

We recall some facts about these groups from [Suz-82]. If $n = 2m-1$ is odd the commutator subgroup $\Omega(2m+1, |k_{\mathcal{P}}|)$ is simple and of index 2 in $\mathrm{SO}(2m+1, |k_{\mathcal{P}}|)$.

If $m \geq 2$ and $n = 2m-2$ even then there are two special orthogonal groups $\mathrm{SO}_{\pm}(2m, |k_{\mathcal{P}}|)$ and we write $\Omega_{\pm}(2m, |k_{\mathcal{P}}|)$ for the commutator subgroup. The center has size at most two and the central factor group is simple for $m \geq 3$. When $m = 2$ we have split and nonsplit versions

$$\mathrm{P}\Omega_+(4, |k_{\mathcal{P}}|) = \mathrm{PSL}_2(|k_{\mathcal{P}}|) \times \mathrm{PSL}_2(|k_{\mathcal{P}}|),$$

$$\mathrm{P}\Omega_-(4, |k_{\mathcal{P}}|) = \mathrm{PSL}_2(|k_{\mathcal{P}}|^2).$$

The following Lemma gives the needed bounds for prime levels.

2.1.1 Lemma *Let φ be a nontrivial representation of $\Lambda_1/\Lambda_1(\mathcal{P})$. Then the dimension of φ is bounded below as $|k_{\mathcal{P}}| \to \infty$ by*

$$\dim \varphi \gg |k_{\mathcal{P}}|^{n-1}. \tag{2.10}$$

We have for the size of the group $\Lambda_1/\Lambda_1(\mathcal{P})$

$$|\Lambda_1/\Lambda_1(\mathcal{P})| \approx |k_{\mathcal{P}}|^{(n+2)(n+1)/2}. \tag{2.11}$$

Proof By the previous discussion, outside finitely many primes we have $\Lambda_1/\Lambda_1(\mathcal{P}) \cong \Omega_{(\pm)}(n+2, |k_{\mathcal{P}}|)$. The possible sizes for this group can be found in [Suz-82]. For $n \neq 2, 4$ this is a perfect central extension of degree at most 2 of a finite Chevalley group, and lower bounds for the dimension of a nontrivial representation of such a group can be found in [SZ-93]. If $n = 2$ then we have $\Lambda_1/\Lambda_1(\mathcal{P})$ a degree 2 perfect central extension of $\mathrm{PSL}_2(|k_{\mathcal{P}}|) \times \mathrm{PSL}_2(|k_{\mathcal{P}}|)$ or $\mathrm{PSL}_2(|k_{\mathcal{P}}|^2)$. At worst we have a faithful irreducible representation of $\mathrm{PSL}_2(|k_{\mathcal{P}}|)$ contained in φ. The needed bound is then well known. Finally if $n = 4$, using the accidental isomorphisms ([Suz-82])

$$\mathrm{P}\Omega_+(6, |k_{\mathcal{P}}|) = \mathrm{PSL}_4(|k_{\mathcal{P}}|), \quad \mathrm{P}\Omega_-(6, |k_{\mathcal{P}}|) = \mathrm{PSU}_4(|k_{\mathcal{P}}|),$$

then in either case there is an associated nontrivial projective representation whose dimension can be bounded by further results tabulated in [SZ-93]. □

2.1.3 Prime power case

In this section we make use of the work of Weisfeiler [Weis-84] to bound the size of the group $\Lambda_1/\Lambda_1(\mathcal{P}^r)$ and the work of Kelmer and Silberman [KeSi-10] along with some improvements to bound the dimension of new representations. As we have $\Lambda_1/\Lambda_1(\mathcal{P}^r) = \ker SN_{\mathcal{P}^r}$ any new representation lifts to a nontrivial representation of the Spin group which we denote $H(\mathcal{O}/\mathcal{P}^r)$. The level structures are such that the lift is a new representation of $H(\mathcal{O}/\mathcal{P}^r)$.

For $i > 0$ let $H(\mathcal{P}^i)$ denote the kernel of the reduction map $H(\mathcal{O}/\mathcal{P}^r) \to H(\mathcal{O}/\mathcal{P}^i)$, or in other words the congruence subgroup of $H(\mathcal{O}/\mathcal{P}^r)$ of level $\mathcal{P}^i$. Let L denote the Lie algebra of H. We will use the following Lemma of Weisfeiler.

2.1.2 Lemma ([Weis-84] Lemma 5.2)

1. *For $i > 0$ the $H(k_\mathcal{P})$ module $H(\mathcal{P}^i)/H(\mathcal{P}^{i+1})$ is isomorphic to $L(k_\mathcal{P}) \otimes \mathcal{P}^i/\mathcal{P}^{i+1}$, where the action on the first factor is by* Ad *and the action on the second factor is trivial.*

2. *The map $(x, y) \mapsto [x, y]$ maps $H(\mathcal{P}^i) \times H(\mathcal{P}^j)$ into $H(P^{i+j})$ and descends to a map*

$$H(\mathcal{P}^i)/H(\mathcal{P}^{i+1}) \times H(\mathcal{P}^j)/H(\mathcal{P}^{j+1}) \to H(\mathcal{P}^{i+j})/H(\mathcal{P}^{i+j+1}).$$

   This map is given explicitly by

$$[x \otimes r, y \otimes s] = [x, y] \otimes rs$$

   when $H(\mathcal{P}^i)/H(\mathcal{P}^{i+1})$ is viewed as $L(k_\mathcal{P}) \otimes \mathcal{P}^i/\mathcal{P}^{i+1}$ and similarly for j, $j + i$.

3. *If $[L(k_\mathcal{P}), L(k_\mathcal{P})] = L(k_\mathcal{P})$ then $[H(\mathcal{P}), H(\mathcal{P})] = H(\mathcal{P}^2)$.*

Immediately it follows that $H(\mathcal{P}^i)$ is abelian for $i \geq r/2$ in light of $[H(\mathcal{P}^i), H(\mathcal{P}^i)] \subseteq H(\mathcal{P}^{2i}) = \{1\}$. Let $k = [r/2]$ be the integral part of $r/2$. Suppose that ρ is a new representation of $H(\mathcal{O}/\mathcal{P}^r)$, i.e. the restriction $\mathrm{Res}^{H(\mathcal{O}/\mathcal{P}^r)}_{H(\mathcal{P}^{r-1})}\rho$ is not trivial. The work of Kelmer and Silberman [KeSi-10] can be paraphrased[2]

[2]We make a slight improvement here by noting that the proof of [KeSi-10, Proposition 4.4] goes through when their e is $n - 1$ for $n = 2, 3, 4$ in our indexing of n.

2.1.3 Lemma (Kelmer, Silberman) *There is a character χ of $H(\mathcal{P}^{r-k})$ appearing in $\mathrm{Res}^{H(\mathcal{O}/\mathcal{P}^r)}_{H(\mathcal{P}^{r-k})}\rho$ which has nontrivial restriction to $H(\mathcal{P}^{r-1})$. Moreover for the orbit of χ under the co-Adjoint action of $H(\mathcal{O}/\mathcal{P}^r)$ we have*

$$M \equiv |\mathrm{Orbit}_{(H(\mathcal{O}/\mathcal{P}^r),\mathrm{co-Ad})}(\chi)| \gg \begin{cases} |k_{\mathcal{P}}|^{r(n-1)} & \text{if } r = 2k \text{ even} \\ |k_{\mathcal{P}}|^{(r-1)(n-1)} & \text{if } r = 2k+1 \text{ odd} \end{cases} \tag{2.12}$$

and there is an immediate bound below for $\dim \rho \geq M$.

For $r = 2k$ this is the result which we will use. Assume now that $r = 2k + 1$ and we will look for the natural strengthening

$$\dim \rho \gg |k_{\mathcal{P}}|^{r(n-1)}. \tag{2.13}$$

As in Lemma 2.1.3 take a character χ which appears in $\mathrm{Res}^{H(\mathcal{O}/\mathcal{P}^r)}_{H(\mathcal{P}^{r-k})}\rho$ and such that $\mathrm{Res}^{H(\mathcal{P}^{r-k})}_{H(\mathcal{P}^{r-1})}\chi$ is nontrivial. We recall some of the ingredients of the proof for our own use. The co-Adjoint action of $H(\mathcal{O}/\mathcal{P}^r)$ on the unitary dual $\widehat{H(\mathcal{O}/\mathcal{P}^{r-k})}$ descends to an action of $H(\mathcal{O}/\mathcal{P}^k)$. Then $\widehat{H(\mathcal{O}/\mathcal{P}^{r-k})}$ is isomorphic to $L(\mathcal{O}/\mathcal{P}^k)$ via a map which intertwines the co-Adjoint and adjoint actions. Under this fixed isomorphism χ is identified with an element $X \in L(\mathcal{O}/\mathcal{P}^k)$ such that $X \neq 0 \mod \mathcal{P}$. An orbit-stabilizer argument then provides enough characters via the Ad-orbit of X. The bound on the size of the stabilizer is obtained by induction and at each stage the nonzero reduction $X_{\mathcal{P}} \in L(k_{\mathcal{P}})$ of X modulo $\mathcal{P}$ is all the data which is needed. For an explicit formulation of the connection between $X_{\mathcal{P}}$ and $\chi_0 \equiv \mathrm{Res}^{H(\mathcal{P}^{r-k})}_{H(\mathcal{P}^{r-1})}\chi$ let Tr denote the Galois trace $\mathrm{Tr} : k_{\mathcal{P}} \to \mathbb{F}_p$ where $|k_{\mathcal{P}}| = p^f$ for some f. Then after the identification $H(\mathcal{P}^{r-1}) \cong L(k_{\mathcal{P}})$ we have for $Z \in L(k_{\mathcal{P}})$

$$\chi(Z) = \chi_0(Z) = \exp\left(\frac{2\pi i \mathrm{Tr}(B(X_{\mathcal{P}}, Z))}{p}\right), \tag{2.14}$$

where B denotes the nondegenerate Killing form on $L(k_{\mathcal{P}})$.

Let V_χ be the subspace of V (the vector space associated to ρ) upon which $H(\mathcal{P}^{r-k})$ acts by χ. As all the $V_{\chi'}$ for χ' in the co-Adjoint orbit are isomorphic and orthogonal, if we can prove the dimension of V_χ is large we get a bound

$$\dim V = \dim \rho \gg \dim V_\chi \, |\mathrm{Orbit}_{(H(\mathcal{O}/\mathcal{P}^r),\mathrm{co-Ad})}(\chi)|. \tag{2.15}$$

In the next Lemma we utilize a better bound on $\dim V_\chi$ to get an improvement for r odd on the dimension bound in Lemma 2.1.3 (which is using the trivial $\dim V_\chi \geq 1$).

2.1.4 Lemma *Let $r = 2k+1 \geq 3$ and ρ a new representation of $H(\mathcal{O}/\mathcal{P}^r)$. Then*

$$\dim \rho \gg |k_{\mathcal{P}}|^{r(n-1)} \tag{2.16}$$

with implied constant uniform through r and $\mathcal{P}$.

Proof Take χ, V_χ as before. We avoid all finitely many primes $\mathcal{P}$ where the characteristic p of $k_{\mathcal{P}}$ is ramified in F. By Lemma 2.1.2 $H(\mathcal{P}^{r-k})$ is in the center of $H(\mathcal{P}^k) = H(\mathcal{P}^{r-k-1})$ and so $H(\mathcal{P}^k)$ preserves V_χ. Let ϕ_0 denote the subrepresentation of $\mathrm{Res}^{H(\mathcal{O}/\mathcal{P}^r)}_{H(\mathcal{P}^k)}\rho$ corresponding to V_χ and choose some irreducible representation (ϕ, W) of $H(\mathcal{P}^k)$ appearing in (ϕ_0, V_χ). Our trick is to consider the $\mathrm{Hom}(W,W) \cong \bar{\phi}\otimes\phi$ representation of $H(\mathcal{P}^k)$. As $H(\mathcal{P}^{r-k})$ acts as an irreducible character on W it acts trivially on $\mathrm{Hom}(W,W)$ and so $\bar{\phi}\otimes\phi$ factors through a representation of the abelian group $H(\mathcal{P}^k)/H(\mathcal{P}^{r-k})$ which is isomorphic to $L(k_{\mathcal{P}})$ by Lemma 2.1.2. Then $\bar{\phi}\otimes\phi$ splits as a direct sum of irreducible characters of $L(k_{\mathcal{P}})$

$$\bar{\phi}\otimes\phi \cong \bigoplus_{i=1}^{(\dim W)^2} \theta_i. \tag{2.17}$$

Similarly to before we can write for each i and $Z \in L(k_{\mathcal{P}})$

$$\theta_i(Z) = \exp\left(\frac{2\pi i \mathrm{Tr}(B(Y_i, Z))}{p}\right) \tag{2.18}$$

for uniquely determined $Y_i \in L(k_{\mathcal{P}})$. Let $U = \langle Y_1, \ldots, Y_{(\dim W)^2}\rangle$ be the $\mathbb{F}_p$ vector subspace of $L(k_{\mathcal{P}})$ spanned by the Y_i. In fact, via composition in $\mathrm{Hom}(W, W)$ any element of this space is one of the Y_i as the θ_i form an abelian group. Let $U^{\perp}$ be the orthogonal subspace to U with respect to $\mathrm{Tr}(B(\bullet, \bullet))$. For $v \in U^{\perp}$ we have $\theta_i(v) = 1$ for all i and therefore $\bar{\phi} \otimes \phi(\tilde{v}) = \mathrm{Id}_{\mathrm{Hom}(W,W)}$, where $\tilde{v}$ denotes a lift of v in $H(\mathcal{P}^k)$. The group G_v generated by $\tilde{v}$ and all of $H(\mathcal{P}^{r-k})$ does not depend on the lift and is an abelian group by Lemma 2.1.2. Therefore $\mathrm{Res}^{H(\mathcal{P}^k)}_{G_v}\phi$ splits as a direct sum of irreducible characters. As $\bar{\phi} \otimes \phi(\tilde{v})$ is trivial all of these characters must be the same and therefore $\phi(\tilde{v})$ is a scalar multiple of the identity on W. For any $\tilde{z} \in H(\mathcal{P}^k)$ reducing mod $H(\mathcal{P}^{r-k})$ to $z \in L(k_{\mathcal{P}})$ we have for the image of the group commutator

$$1 = \phi([\tilde{v}, \tilde{z}]_{\mathbf{Grp}}) = \phi([v, z]_{\mathbf{Lie}}), \tag{2.19}$$

where $[v, z]_{\mathbf{Lie}}$ is viewed as an element of $H(\mathcal{P}^{r-1})$ by Lemma 2.1.2. Moreover

$$\phi([\tilde{v}, \tilde{z}]) = \chi_0([v, z]) = \exp\left(\frac{2\pi i \mathrm{Tr}(B(X_{\mathcal{P}}, [v, z]))}{p}\right)$$

in the notation from our previous discussion so in particular

$$\mathrm{Tr}(B(X_{\mathcal{P}}, [v, z])) = 0 \tag{2.20}$$

for all $z \in L(k_{\mathcal{P}})$, or what is the same by ad-invariance of the Killing form

$$\mathrm{Tr}(B([v, X_{\mathcal{P}}], z)) = 0. \tag{2.21}$$

By nondegeneracy it follows that $[v, X_{\mathcal{P}}] = 0$, i.e. v is in the centralizer $C_{L(k_{\mathcal{P}})}(X_{\mathcal{P}})$. As in [KeSi-10] we have[3]

$$\dim_{k_{\mathcal{P}}} C_{L(k_{\mathcal{P}})}(X_{\mathcal{P}}) \leq \dim_{k_{\mathcal{P}}} L(k_{\mathcal{P}}) - 2(n-1) \tag{2.22}$$

and as we have shown $U^{\perp} \subseteq C_{L(k_{\mathcal{P}})}(X_{\mathcal{P}})$ it follows that

$$\begin{aligned} \dim_{\mathbb{F}_p} U^{\perp} \leq \dim_{\mathbb{F}_p} C_{L(k_{\mathcal{P}})}(X_{\mathcal{P}}) &= f \dim_{k_{\mathcal{P}}} C_{L(k_{\mathcal{P}})}(X_{\mathcal{P}}) \\ &\leq f(\dim_{k_{\mathcal{P}}} L(k_{\mathcal{P}}) - 2(n-1)) \\ &= \dim_{\mathbb{F}_p} L(k_{\mathcal{P}}) - 2f(n-1) \end{aligned}$$

where f is the inertia degree. Then immediately $\dim_{\mathbb{F}_p} U \geq 2f(n-1)$ so that

$$\dim W \geq \sqrt{\#\text{ of } Y_i} = \sqrt{|\mathbb{F}_p|^{\dim_{\mathbb{F}_p} U}} \geq \sqrt{|\mathbb{F}_p|^{2f(n-1)}} = |k_{\mathcal{P}}|^{n-1}. \tag{2.23}$$

Finally W is a subspace of V_{χ} and as we remarked before we have the bound

$$\dim \rho \geq \dim V_{\chi} |\mathrm{Orbit}_{(H(\mathcal{O}/\mathcal{P}^r),\mathrm{co-Ad})}(\chi)| \geq \dim W |\mathrm{Orbit}_{(H(\mathcal{O}/\mathcal{P}^r),\mathrm{co-Ad})}(\chi)|. \tag{2.24}$$

Inserting the bound from Lemma 2.1.3 on the size of the orbit together with the bound in (2.23) gives us the desired result

$$\dim \rho \gg |k_{\mathcal{P}}|^{n-1} |k_{\mathcal{P}}|^{(r-1)(n-1)} = |k_{\mathcal{P}}|^{r(n-1)}. \tag{2.25}$$

□

We remark that the idea of this proof should hold in the more general setting of Lemma 2.1.2, provided centralizer bounds analogous to (2.22) can be obtained.

[3] This inequality is the key to the inductive step of the proof of Lemma 2.1.3, so it is fitting that it is also the ingredient here.

2.1.4 General result

Now we have all that is required for the main result of this section.

2.1.5 Lemma *If*

$$\mathcal{I} = \prod_{i=1}^{l} \mathcal{P}_i^{r_i} \tag{2.26}$$

is the prime factorization of $\mathcal{I}$ an ideal in $\mathcal{O}_F$, then for the size of the group $\Lambda_1/\Lambda_1(\mathcal{I})$ we have

$$|\Lambda_1/\Lambda_1(\mathcal{I})| \approx \prod_{i=1}^{l} |k_{\mathcal{P}_i}|^{r_i(n+1)(n+2)/2} = |\mathcal{O}_F/\mathcal{I}|^{(n+1)(n+2)/2}. \tag{2.27}$$

Any representation ρ of $\Lambda_1/\Lambda_1(\mathcal{I})$ of level $\mathcal{I}$ has dimension

$$\dim \rho \gg \prod_{i=1}^{l} |k_{\mathcal{P}_i}|^{r_i(n-1)} = |\mathcal{O}_F/\mathcal{I}|^{n-1}. \tag{2.28}$$

Proof For the size of the group it is sufficient to give the size of each of the $\Lambda_1/\Lambda_1(\mathcal{P}_i^{r_i})$ by (2.7). This is obtained from the bound at prime level in Lemma 2.1.1 together with Lemma 2.1.2 which gives the size of the factor $H(\mathcal{P}_i^j)/H(\mathcal{P}_i^{j+1})$ as $|k_{\mathcal{P}_i}|^{(n+2)(n+1)/2} = |L(k_{\mathcal{P}_i})|$. The bound on the dimension of ρ follows from the discussion leading up to (2.8) together with the bounds obtained at prime level in Lemma 2.1.1, at even power of prime level in Lemma 2.1.3 and at odd power of prime level in Lemma 2.1.4. □

2.2 Eigenfunction Estimates

2.2.1 The geometry near infinity

Recall that we are considering geometrically finite, torsion free and orientation preserving subgroups $\Lambda, \Lambda(\mathcal{I})$ of $\mathrm{Isom}^+(\mathbb{H}^{n+1})$ (here we replace Λ with the Λ_1 of

the previous section if necessary). The geometry of the quotients $X(\mathcal{I}) = \mathbb{H}^{n+1}/\Lambda(\mathcal{I})$ is well understood, although less is usually said about the nature of the covering maps $\pi_{\mathcal{I}} : X(\mathcal{I}) \to X$. We aim in this section to describe the geometry of the covering maps near infinity. A good description of the geometry at a fixed level can be found in the paper of Mazzeo and Phillips [MP-90]. We follow the notation of [Gu-09], where Guillarmou considers slightly less general spaces (with finite holonomy in the cusps) but provides very useful analytic lemmas which can be extended to our case without difficulty.

The space $\mathbb{H}^{n+1}$ has a natural compactification by adding a sphere at infinity S^n_∞. The action of Λ on $\mathbb{H}^{n+1}$ extends to the sphere. If o is a point in $\mathbb{H}^{n+1}$ then the orbit Λo accumulates on a subset of the boundary denoted $L(\Lambda)$, the limit set of Λ (that this set is independent of o is a general property of nonpositive curvature). The complement $S^n_\infty - L(\Lambda)$ is called the domain of discontinuity. It follows that Λ acts discretely and properly discontinuously on the domain of discontinuity. As Λ is geometrically finite it has a Dirichlet fundamental domain $\mathcal{F}$ in $\mathbb{H}^{n+1}$ which is finitely faced by totally geodesic hypersurfaces. If Λ is infinite index in Γ then this fundamental domain necessarily extends to the boundary, and the bounding hypersurfaces meet the boundary in subsets of $n-1$ dimensional spheres. Similar statements hold for $\Lambda(\mathcal{I})$, in particular a fundamental domain $\mathcal{F}_{\mathcal{I}}$ for $\Lambda(\mathcal{I})$ is paved by images of $\mathcal{F}$ under coset representatives of $\Lambda/\Lambda(\mathcal{I})$.

The elements of Γ can be classified by their fixed points on the boundary. Either γ fixes two points on the boundary and the geodesic between these two points,

in which case it is called *hyperbolic*, or γ has one fixed point and acts by Euclidean motions on horospheres tangent to this point. In the latter case γ is called *parabolic*.

Away from parabolic fixed points, the region where $\mathcal{F}$ meets the boundary can be covered by finitely many charts isometric to regions

$$M_r = \{(x, y) \in (0, \infty) \times \mathbb{R}^n \; ; \; x^2 + |y|^2 < 1\}, \tag{2.29}$$
$$g_r = x^{-2}(dx^2 + dy^2),$$

which we call regular neighbourhoods. These can be chosen sufficiently small so that they project isometrically to the quotient X. The pulled back charts cover a corresponding region in $X(\mathcal{I})$. The remaining neighbourhoods are near parabolic fixed points. Suppose that ∞ in the upper half space model is the fixed point of a parabolic element, by conjugating if necessary. Let Λ_∞ consist of all elements of Λ with this fixed point. This group is purely parabolic by discreteness and can be thought of as acting as Euclidean isometries on any horosphere, which is isomorphic to $\mathbb{R}^n$. Some of the facts which follow come from the theory of Bieberbach groups for which the reader can consult the notes of Thurston [T-97].

Following [MP-90], the group Λ_∞ contains a maximal normal free abelian subgroup Λ_a of finite index. We define the rank k of the cusp at ∞ to be the rank of Λ_a. There exists a maximal affine subspace $\mathbb{R}^k \subset \mathbb{R}^n$ fixed by Λ_∞. The subgroup Λ_a acts on this space by translations and the quotient $F_k \equiv \mathbb{R}^n/\Lambda_\infty$ is the total space of a flat vector bundle of rank $n - k$ over a compact flat base manifold B_k. This B_k can realized concretely as $\mathbb{R}^k/\Gamma_\infty$. It is covered by a flat k-torus $T^k = \mathbb{R}^k/\Lambda_a$ by the usual Galois correspondence, the covering map coming from the map of flat vector bundles

$\tilde{F}_k \equiv \mathbb{R}^n/\Lambda_a \to \mathbb{R}^n/\Lambda_\infty$ restricted to the zero section. We use y for a local coordinate in the fibre coming from $\mathbb{R}^{n-k}$, and z for a coordinate on B_k coming from the covering $\mathbb{R}^k \to \mathbb{R}^k/\Lambda_\infty$. To cover the regions at infinity in X which come from parabolic fixed points we can use charts isometric to rank k cuspidal neighbourhoods

$$M_k = \{(x,[y,z]) \in (0,\infty) \times F_k \,;\, x^2 + |y|^2 > 1\}, \tag{2.30}$$
$$g_k = x^{-2}(dx^2 + dy^2 + dz^2),$$

where we are writing $[y,z]$ for local trivializing coordinates. The quadratic differential dz^2 refers to a flat metric on B_k. Note also that $|y|^2$ is a well defined function, as changing trivialization affects an orthogonal transformation on y and similarly dy^2 is defined independently of trivialization. In general there are finitely many cusps of each rank but to simplify the discussion we assume that there is only one neighbourhood of each cuspidal type M_k. We drop the isometries which identify the sets in X with the model neighbourhoods and think of the M_r, M_k as sets in X. It can be arranged so that all the cuspidal neighbourhoods are disjoint.

We consider now the covering maps $\pi_{\mathcal{I}} : X_{\mathcal{I}} \to X$. If q is a parabolic fixed point of Λ with stabilizer Λ_q then $\Lambda_q \cap \Gamma(\mathcal{I})$ also fixes this point and is the stabilizer in $\Lambda(\mathcal{I})$, i.e. $\Lambda(\mathcal{I})_q = \Lambda_q \cap \Gamma(\mathcal{I})$. In addition, $\Lambda_a \cap \Gamma(\mathcal{I})$ is the maximal normal free abelian subgroup $\Lambda(\mathcal{I})_a$ in $\Lambda(\mathcal{I})_q$. There is then a map of flat rank $n-k$ vector bundles $F_{k,\mathcal{I}} \equiv \mathbb{R}^n/\Lambda(\mathcal{I})_q \to F_k$ which when restricted to the zero section gives a covering map of flat compact manifolds $B_{k,\mathcal{I}} \equiv \mathbb{R}^k/\Lambda(\mathcal{I})_q \to B_k$. If one considers the images of $\mathcal{F}$ under coset representatives of $\Lambda/\Lambda(\mathcal{I})$ one sees that $\pi_{\mathcal{I}}^{-1}(M_k)$ is isometric to a disjoint

union of isometric cuspidal neighbourhoods

$$\pi_{\mathcal{I}}^{-1}(M_k) \cong \coprod_{i=1}^{m_{k,\mathcal{I}}} M_{k,\mathcal{I}} \tag{2.31}$$

where

$$M_{k,\mathcal{I}} = \{(x,[y,z]) \in (0,\infty) \times F_{k,\mathcal{I}} \,;\, x^2 + |y|^2 > 1\}, \tag{2.32}$$

$$g_k = x^{-2}(dx^2 + dy^2 + dz^2),$$

and as before, y and z refer to trivializing coordinates for the flat vector bundle. If $l_{k,\mathcal{I}}$ is the degree of the covering of base manifolds $B_{k,\mathcal{I}} \to B_k$ then we have $m_{k,\mathcal{I}} l_{k,\mathcal{I}} = |\Lambda/\Lambda(\mathcal{I})|$, the degree of the covering map $\pi_{\mathcal{I}}$. As we will work with the covering tori $T^k_{\mathcal{I}} \equiv \mathbb{R}^k/\Lambda(\mathcal{I})_a \to T_k$ it is salient to note that the obvious diagram of covering maps involving the tori $T^k_{\mathcal{I}}$, T^k and base manifolds $B_{k,\mathcal{I}}$, $B_{\mathcal{I}}$ is commutative. The covering map at M_k can be given explicitly with respect to these neighbourhoods, indeed it is directly induced by the map $F_{k,\mathcal{I}} \to F_k$ so that the x coordinate is preserved. As for the regular neighbourhoods, it can be arranged so that the preimage $\pi_{\mathcal{I}}^{-1}(M_r)$ of each individual M_r is isometric to $|\Lambda/\Lambda(\mathcal{I})|$ disjoint copies of M_r. Then the covering map in each disjoint copy is given by the identity with respect to these charts.

Now we seek compactification coordinates for these charts. The transformation defined locally and which only affects the coordinate in $\mathbb{R}^+$ and in the fibre

$$(x,[y,z]) \mapsto (t,[u,z]) = \left(\frac{x}{x^2+|y|^2}, [\frac{-y}{x^2+|y|^2}, z]\right) \tag{2.33}$$

actually (by flatness) gives a diffeomorphism from (M_k, g_k) to

$$\{(t,[u,z]) \in (0,\infty) \times F_k \,;\, t^2 + |u|^2 < 1\}.$$

The pushed forward metric is

$$t^{-2}(dt^2 + du^2 + (t^2 + |u|^2)^2 dz^2). \tag{2.34}$$

These coordinates allow the charts M_k to be smoothly compactified by adding a $\{t = 0\}$ portion to form $\bar{M}_k$. The charts M_r also naturally compactify by adding a $\{x = 0\}$ part and all these boundary pieces join together to give a smooth boundary $\delta\bar{X}$ to X. The compactification of X is denoted $\bar{X}$, this has the structure of a smooth compact manifold with boundary $\delta\bar{X}$. A similar procedure takes place to smoothly compactify $X(\mathcal{I})$ to $\bar{X}(\mathcal{I})$. For each M_k there is a cusp submanifold isomorphic to B_k of $\delta\bar{X}$ corresponding to the zero section of F_k at $t = 0$. The cusp submanifold of $\delta\bar{X}$ coming from M_k is denoted b_k and we define

$$B \equiv \delta\bar{X} - \coprod_{k=1}^{n} b_k. \tag{2.35}$$

Following the same procedure at level $\mathcal{I}$ we get a regular boundary part $B(\mathcal{I})$ which naturally covers B. A collar neighbourhood of the boundary could now be expected by Milnor's Collar Neighbourhood Lemma [Mi-65]; the following proposition gives us some fine control over the geometry in such a neighbourhood.

2.2.1 Proposition *In some collar neighbourhood $(0, \epsilon)_\rho \times \delta\bar{X}$ of $\delta\bar{X}$ the hyperbolic metric is given by*

$$g = \frac{d\rho^2 + h(\rho)}{\rho^2} \tag{2.36}$$

for some smooth family of symmetric tensors $h(\rho)$ on $\delta\bar{X}$, depending smoothly on ρ, positive for $\rho > 0$ with $h(0) = h_0$ positive on B and satisfying

$$h(\rho) = du^2 + (\rho^2 + |u|^2)^2 dz^2$$

in each $\bar{M}_k$. *Moreover* $\rho = t$ *in* $\bar{M}_k$.

This appears as a discussion in [Gu-09]. It extends to include maximal rank cusps by interpreting the u variable as absent. Also, Guillarmou is only considering the case of rational cusps, but the proof goes through in our slightly more general case as it only relies on a PDE being non-characteristic away from the cusps.

The boundary defining function ρ lifts from X to $X(\mathcal{I})$. By our previous remarks on the nature of the covering maps in local charts, by making ϵ small enough so $\rho^{-1}((0,\epsilon))$ does not escape any of the local charts we can use the same ϵ for all $\mathcal{I}$. The metric provided also lifts to the collar neighbourhood $(0,\epsilon)_{\pi_{\mathcal{I}}^*\rho} \times \delta\bar{X}(\mathcal{I})$. By examining the covering map in the cuspidal regions we get an exact form for the metric at each level $\mathcal{I}$.

2.2.2 Bounds below in the cusps

Here we examine the cuspidal regions near infinity. There the metric is exact and we can get an exact result. In the local $[u,z]$ coordinates corresponding to the boundary part in $\bar{M}_k$, define the product regions $N_k(R) = (0,1/R) \times \{|u|^2 < 3/4\}$, and $N_{k,\mathcal{I}}(R) = \pi_{\mathcal{I}}^{-1}(N_k(R))$. In the maximal rank cusp $N_n(R) = (0,1/R) \times B_n$. Note that by choosing R_0 large enough, we can assume that for all $R > R_0$, $N_k(R) \subset \bar{M}_k$ (hence the corresponding result at level $\mathcal{I}$), and we make this increase for R_0 immediately if necessary, and consider $R > R_0$.

We separate into two cases depending on whether the holonomy representa-

tion

$$h : \pi_1(T^k) \to \mathrm{O}(n-k) \tag{2.37}$$

has finite image. If it does then we can pass to a finite locally isometric covering $f : (\overline{T^k}, f^*dz^2) \to (B_k, dz^2)$ such that f^*F_k is trivial. The torus $\overline{T^k}$ is the quotient $\mathbb{R}^k/\ker(h)$. Moreover the holonomy representation of $T^k_{\mathcal{I}}$ is the restriction to a smaller group so remains finite at all levels. The reader can also note that the finite holonomy property is always satisfied for $n = 2$. We cover finite holonomy in the following Lemma.

2.2.2 Lemma *If the holonomy representation of T^k has finite image then*

$$\Delta_{X(\mathcal{I})}|_{C_0^\infty(N_{k,\mathcal{I}}(R))} \geq n^2/4. \tag{2.38}$$

Proof First we make a lifting argument to simplify the case. Let $\phi \in C_0^\infty(N_{k,\mathcal{I}}(R))$. We lift ϕ to a region covering $N_{k,\mathcal{I}}(R)$

$$\overline{N_{k,\mathcal{I}}(R)} \equiv \{(t,u,z) \in (0,1/R) \times \mathbb{R}^{n-k} \times \overline{T^k_{\mathcal{I}}} \ : \ |u|^2 < 3/4\}$$

equipped with the pulled back metric via $1 \times f$. Then note

$$\langle \Delta_{\overline{N_{k,\mathcal{I}}(R)}}(1\times f)^*\phi, (1\times f)^*\phi\rangle_{\overline{N_{k,\mathcal{I}}(R)}} = |h(\pi_1(T^k_{\mathcal{I}}))| \times \langle \Delta_{N_{k,\mathcal{I}}(R)}\phi, \phi\rangle_{N_{k,\mathcal{I}}(R)},$$

$$\|(1\times f)^*\phi\|^2_{\overline{N_{k,\mathcal{I}}(R)}} = |h(\pi_1(T^k_{\mathcal{I}}))| \times \|\phi\|^2_{N_{k,\mathcal{I}}(R)},$$

and $(1\times f)^*\phi \in C_0^\infty(\overline{N_{k,\mathcal{I}}(R)})$ so it is sufficient to prove the result when the cross section is a trivial bundle over a torus. Then assume that $F_{k,\mathcal{I}} \cong \mathbb{R}^{n-k} \times T^k_{\mathcal{I}}$, so that

$$N_{k,\mathcal{I}}(R) = \{(t,u,z) \in (0,1/R) \times \mathbb{R}^{n-k} \times \overline{T^k_{\mathcal{I}}} \ : \ |u|^2 < 3/4\}.$$

Now we introduce the change of coordinates $t = e^{-\tau}$, and then conjugate the Laplacian by the function $|g|^{1/4} = e^{n\tau/2}(e^{-2\tau} + |u|^2)^{k/2}$. Clearly $\phi \mapsto |g|^{1/4}\phi$ preserves $C_0^\infty(N_{k,\mathcal{I}}(R))$. The conjugated Laplacian acts on the L^2 space defined by the volume element $d\tau \wedge du \wedge dz$. One has for $L = |g|^{1/4}\Delta|g|^{-1/4}$ in coordinates (τ, u, z)

$$L = -\partial_\tau^2 + e^{-2\tau}\Delta_u^{\text{Euclidean}} + \frac{e^{-2\tau}}{(e^{-2\tau} + |u|^2)^2}\Delta_z^{\text{FlatToroidal}} + \frac{n^2}{4}. \tag{2.39}$$

The main part of this calculation appears in [Gu-09, eq. 5.2]. This completes the proof as the first 3 terms can be easily verified nonnegative on $C_0^\infty(N_{k,\mathcal{I}}(R))$ with the product measure. □

As remarked before the reader can skip the next Lemma if they are interested only in the case $n = 2$. For the remaining cases the idea is that functions on a flat bundle with nondiscrete holonomy are equivalent to functions on a bundle with discrete holonomy where the previous Lemma can be applied. Moreover we can find finite holonomy bundles 'arbitrarily close' to the original, and the action of the Laplacian is continuous in some sense with respect to this approximation. This perturbation argument is due to Mazzeo and Phillips as they use it in [MP-90, Lemma 5.12].

2.2.3 Lemma *In any cuspidal neighbourhood $N_{k,\mathcal{I}}(R)$,*

$$\Delta_{X(\mathcal{I})}|_{C_0^\infty(N_{k,\mathcal{I}}(R))} \geq n^2/4. \tag{2.40}$$

Proof We assume now that the holonomy representation has infinite image, or we are done by Lemma 2.2.2. We drop the $\mathcal{I}$ dependence which does not matter as we only deal with one level at a time. By the same lifting argument as before it is sufficient

to consider the case when $B_k = T^k$ so the cross section is a flat bundle F_k over a torus. Then conjugating the parabolic fixed point to ∞ in the upper half space model $\{(x, u, z) \in \mathbb{R}^+ \times \mathbb{R}^{n-k} \times \mathbb{R}^k\}$ we have $\Lambda_\infty = \Lambda_a = \pi_1(T^k)$ a free abelian group of rank k and preserving the Euclidean horosphere $\{x = 1\}$ which we take as a cover. The image of the holonomy representation is commutative so each element in the image has the same invariant subspaces. For simplicity we proceed as though $n = 3$ and $k = 1$, so that $\pi_1(T^1)$ is infinite cyclic with generator γ. By hypothesis (and further conjugating if necessary) γ acts on the plane $\{(x, u, z) \in \mathbb{R}^+ \times \mathbb{R}^2 \times \mathbb{R} : x = 1\}$ by

$$\gamma : (u, z) \mapsto \left(\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} u, z + 1 \right)$$

with θ irrational. We consider the perturbing map

$$p_\eta : (u, z) \mapsto \left(\begin{pmatrix} \cos(\eta z) & \sin(\eta z) \\ -\sin(\eta z) & \cos(\eta z) \end{pmatrix} u, z \right)$$

where $\eta + \theta$ is rational and η will be chosen arbitrarily small in what follows. We compute

$$p_\eta \gamma p_\eta^{-1}(u, z) = \left(\begin{pmatrix} \cos(\theta + \eta) & \sin(\theta + \eta) \\ -\sin(\theta + \eta) & \cos(\theta + \eta) \end{pmatrix} u, z + 1 \right).$$

This implies p_η descends to a bundle isomorphism from $F_k \to F_k^\eta$ where F_k^η is the total space of a bundle over T^1 with rational holonomy. It also induces a map (also denoted p_η) to a new cuspidal region $N_k^\eta(R)$ in the obvious way. Now take $\phi \in C_0^\infty(N_k(R))$.

It is not hard to show[4] that

$$\langle \Delta_{N_k^\eta(R)} p_{\eta *}\phi, p_{\eta *}\phi\rangle_{L^2(N_k^\eta(R))} = \langle p_\eta^* \Delta_{N_k^\eta(R)} p_{\eta *}\phi, \phi\rangle_{L^2(N_k(R))} \to \langle \Delta_{N_k(R)}\phi,\phi\rangle_{L^2(N_k(R))} \tag{2.41}$$

as $\eta \to 0$. We have also

$$\langle \Delta_{N_k^\eta(R)} p_{\eta *}\phi, p_{\eta *}\phi\rangle_{L^2(N_k^\eta(R))} \geq n^2/4 \|p_{\eta *}\phi\|^2_{L^2(N_k^\eta)(R)} = n^2/4\|\phi\|^2_{L^2(N_k(R))}$$

by Lemma 2.2.2. Then choosing a sequence of $\eta_i \to 0$ with $\theta + \eta_i$ rational in equation (2.41) we have

$$\langle \Delta_{N_k(R)}\phi,\phi\rangle_{L^2(N_k(R))} \geq n^2/4\|\phi\|^2_{L^2(N_k(R))}$$

which is the required bound below. The same idea works in the general case with more 'perturbation directions'. □

2.2.3 Bounds below in regular neighbourhoods

Let $B_0 \subset \delta\bar{X}$ be the compact manifold with boundary

$$\begin{aligned} B_0 &= \delta\bar{X} - \bigcup_{k=1}^{n-1} \{(u,z) \in \delta\bar{M}_k : |u|^2 < 1/4\}, \\ \delta B_0 &= \bigcup_{k=1}^{n-1} \{(u,z) \in \delta\bar{M}_k : |u|^2 = 1/4\}. \end{aligned} \tag{2.42}$$

By Proposition 2.2.1, after the change of coordinates $\rho = e^{-\tau}$ the hyperbolic metric on $\pi_{\mathcal{I}}^{-1}((0,1/R)\times B_0)$ is of the form

$$g = d\tau^2 + e^{2\tau}\pi_{\mathcal{I}}^*\gamma(\tau).$$

[4] Something very similar appears in the proof of [MP-90, Lemma 5.12].

Here $\gamma(\tau) = h(e^{-\tau})$ with h as in Proposition 2.2.1. The Laplacian on $\pi_{\mathcal{I}}^{-1}((\log R, \infty)_\tau \times B_0)$ takes the form (as in [Gu-09, eq. 5.1])

$$\Delta_{X(\mathcal{I})} = -\partial_\tau^2 - n\partial_\tau - \frac{1}{2}\pi_{\mathcal{I}}^* \left(\mathrm{Tr}(\gamma^{-1}(\tau).\partial_\tau\gamma)\right) \partial_\tau + e^{-2\tau}\Delta_{\pi_{\mathcal{I}}^*\gamma(\tau)}. \tag{2.43}$$

Our next Lemma prepares the way by giving the Laplacian on $\pi_{\mathcal{I}}^{-1}((\log R, \infty)_\tau \times B_0)$ as a 'polynomial' in ∂_τ and $\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}$.

2.2.4 Lemma *The Laplacian on $\pi_{\mathcal{I}}^{-1}((\log R, \infty)_\tau \times B_0)$ can be written*

$$\begin{aligned}\Delta_{X(\mathcal{I})} = -\partial_\tau^2 - n\partial_\tau + e^{-2\tau}\Delta_{\pi_{\mathcal{I}}^*\gamma(0)} \\ e^{-\tau}(\pi_{\mathcal{I}}^* f)\partial_\tau + e^{-3\tau}\left(\pi_{\mathcal{I}}^* a.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} + \mathrm{div}_{\pi_{\mathcal{I}}^*\gamma(0)}.\pi_{\mathcal{I}}^* b.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}\right),\end{aligned} \tag{2.44}$$

where f (resp. a, resp. b) is a smooth bounded function (resp. one form, resp. endomorphism of the tangent bundle) on $(\log R, \infty)_\tau \times B_0$.

Proof By the smoothness of the family of metrics $h(\rho)$, we can write

$$\gamma^{-1}(\tau) = (1 + e^{-\tau}b)\gamma^{-1}(0), \tag{2.45}$$

$$|\gamma(\tau)|^{1/2} = (1 + e^{-\tau}c)|\gamma(0)|^{1/2}, \tag{2.46}$$

$$|\gamma(\tau)|^{-1/2} = (1 + e^{-\tau}\tilde{c})|\gamma(0)|^{-1/2}, \tag{2.47}$$

with (by multiplying (2.46) and (2.47))

$$c + \tilde{c} + e^{-\tau}c\tilde{c} \equiv 0. \tag{2.48}$$

The quantities b, c and $\tilde{c}$ are smooth bounded 2-tensors and functions respectively on $(\log R, \infty)_\tau \times B_0$. It follows from the compactness of B_0 that any fixed finite number of

derivatives of b, c and $\tilde{c}$ are smooth and bounded on $(\log R, \infty)_\tau \times B_0$. The analogous statements hold for $\pi_{\mathcal{I}}^* \gamma(\tau)$ by replacing b, c and $\tilde{c}$ with their lifts.

We calculate that, using (2.46) and (2.47) and writing d for the exterior derivative on B_0,

$$\begin{aligned} \mathrm{div}_{\gamma(\tau)} &= \mathrm{div}_{\gamma(0)} + e^{-\tau}(1 + e^{-\tau}\tilde{c}) d(c) \\ &\quad + e^{-\tau}(c + \tilde{c} + e^{-\tau} c\tilde{c}) \mathrm{div}_{\gamma(0)} \\ &= \mathrm{div}_{\gamma(0)} + e^{-\tau}\omega, \end{aligned} \tag{2.49}$$

where the term on the second line vanished due to (2.48) and ω is a smooth bounded one form on $(\log R, \infty)_\tau \times B_0$ (we use boundedness of derivatives of c here). By (2.45),

$$\nabla_{\gamma(\tau)} = \nabla_{\gamma(0)} + e^{-\tau} b . \nabla_{\gamma(0)}. \tag{2.50}$$

Now using that

$$\Delta_{\gamma(\tau)} = \mathrm{div}_{\gamma(\tau)} . \nabla_{\gamma(\tau)} \tag{2.51}$$

our previous formulae (2.49) and (2.50) give

$$\Delta_{\gamma(\tau)} = \left(\mathrm{div}_{\gamma(0)} + e^{-\tau}\omega\right) . \left(\nabla_{\gamma(0)} + e^{-\tau} b . \nabla_{\gamma(0)}\right) \tag{2.52}$$

$$= \Delta_{\gamma(0)} + e^{-\tau}\omega . \left(1 + e^{-\tau} b\right) . \nabla_{\gamma(0)} + e^{-\tau} \mathrm{div}_{\gamma(0)} . b . \nabla_{\gamma(0)} \tag{2.53}$$

$$= \Delta_{\gamma(0)} + e^{-\tau}\left(a . \nabla_{\gamma(0)} + \mathrm{div}_{\gamma(0)} . b . \nabla_{\gamma(0)}\right), \tag{2.54}$$

where a (resp. b) is a smooth bounded one form (resp. endomorphism of the tangent bundle) on $(\log R, \infty)_\tau \times B_0$. To get the analogous result at level $\mathcal{I}$ we can repeat the argument and note that all the quantities which appear are the lifts of their

counterparts in the previous discussion. This yields

$$\Delta_{\pi_{\mathcal{I}}^*\gamma(\tau)} = \Delta_{\pi_{\mathcal{I}}^*\gamma(0)} + e^{-\tau}\left(\pi_{\mathcal{I}}^* a.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} + \mathrm{div}_{\pi_{\mathcal{I}}^*\gamma(0)}.\pi_{\mathcal{I}}^* b.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}\right). \tag{2.55}$$

Similar arguments show that

$$-\frac{1}{2}\pi_{\mathcal{I}}^*\left(\mathrm{Tr}(\gamma^{-1}(\tau).\partial_\tau\gamma)\right) = e^{-\tau}\pi_{\mathcal{I}}^* f \tag{2.56}$$

where f is a smooth bounded function. Substituting (2.55) and (2.56) into (2.43) gives the desired expression (2.44). □

To proceed, we conjugate the Laplacian by the function

$$G_{\mathcal{I}} \equiv \pi_{\mathcal{I}}^*(|g|/|\gamma(0)|)^{1/4}$$

so as to act on the product space associated to the volume element $d\tau \wedge \mu_{\pi_{\mathcal{I}}^*\gamma(0)}$. The conjugated Laplacian

$$L(\mathcal{I};R) : C_0^\infty(\pi_{\mathcal{I}}^{-1}((\log R,\infty)\times B_0)) \to C_0^\infty(\pi_{\mathcal{I}}^{-1}((\log R,\infty)\times B_0))$$

$$L(\mathcal{I};R) \equiv G_{\mathcal{I}}\Delta_{X(\mathcal{I})}G_{\mathcal{I}}^{-1}$$

takes the form

$$L(\mathcal{I};R) = L_0(\mathcal{I};R) + E(\mathcal{I};R) \tag{2.57}$$

where

$$L_0(\mathcal{I};R) = -\partial_\tau^2 + e^{-2\tau}\Delta_{\pi_{\mathcal{I}}^*\gamma(0)} + n^2/4. \tag{2.58}$$

The operator $\Delta_{\gamma(0)}$ refers to the Laplacian on $(B_0,\gamma(0))$ with Dirichlet boundary conditions. In the next Lemma we compute the error term $E(\mathcal{I};R)$.

2.2.5 Lemma *In the region $\pi_{\mathcal{I}}^{-1}((\log R, \infty)_\tau \times B_0)$ the error term $E(\mathcal{I}; R)$ is given by*

$$E(\mathcal{I}; R) = e^{-\tau} A_{\mathcal{I}} \partial_\tau + e^{-3\tau} B_{\mathcal{I}} . \nabla_{\pi_{\mathcal{I}}^* \gamma(0)} + e^{-3\tau} \mathrm{div}_{\pi_{\mathcal{I}}^* \gamma(0)} . C_{\mathcal{I}} . \nabla_{\pi_{\mathcal{I}}^* \gamma(0)} + e^{-\tau} D_{\mathcal{I}} \quad (2.59)$$

where $A_{\mathcal{I}}$ and $D_{\mathcal{I}}$ are $\mathcal{I}$-uniformly bounded functions, $B_{\mathcal{I}}$ is an $\mathcal{I}$-uniformly bounded one form, and $C_{\mathcal{I}}$ is an $\mathcal{I}$-uniformly bounded endomorphism of the tangent bundle.

Proof Our starting point is equation (2.44). We will compute $E((1); R)$ at full level and compare the calculation to that of general level. Throughout this calculation we will accumulate error terms, we will always use E_i to denote a smooth bounded function on $(\log R, \infty) \times B_0$ and Ω_i to denote a smooth bounded one form on $(\log R, \infty) \times B_0$. We write $G = G_{(1)}$ and note

$$G = e^{\frac{n}{2}\tau}(1 + e^{-\tau} J), \quad (2.60)$$

and

$$G^{-1} = e^{-\frac{n}{2}\tau}(1 + e^{-\tau} J'), \quad (2.61)$$

where $J, J' \in C^\infty((\log R, \infty) \times B_0)$ have bounded derivatives. We will conjugate the terms in equation (2.44) in turn. Firstly we calculate

$$G\left(-\partial_\tau^2 - n\partial_\tau\right) G^{-1} = -\partial_\tau^2 - \left(n + 2G\partial_\tau(G^{-1})\right)\partial_\tau - G\left(\partial_\tau^2(G^{-1}) + n\partial_\tau(G^{-1})\right). \quad (2.62)$$

From (2.61) we can write

$$\partial_\tau(G^{-1}) = -\frac{n}{2} G^{-1} + e^{-(\frac{n}{2}+1)\tau}\left(\partial_\tau J' - J'\right). \quad (2.63)$$

and taking another derivative gives

$$\partial_\tau^2(G^{-1}) = \frac{n^2}{4} G^{-1} + e^{-(\frac{n}{2}+1)\tau}\left(\partial_\tau^2 J' - (n+2)\partial_\tau J' + (n+1)J'\right). \quad (2.64)$$

Using (2.63) and (2.64) in (2.62) gives

$$G\left(-\partial_\tau^2 - n\partial_\tau\right)G^{-1} = -\partial_\tau^2 - 2Ge^{-(\frac{n}{2}+1)}\left(\partial_\tau J' - J'\right)\partial_\tau + \frac{n^2}{4} - Ge^{-(\frac{n}{2}+1)\tau}\left(\partial_\tau^2 J' - 2\partial_\tau J' + J'\right).$$

This can be written

$$G\left(-\partial_\tau^2 - n\partial_\tau\right)G^{-1} = -\partial_\tau^2 + \frac{n^2}{4} + e^{-\tau}E_1\partial_\tau + e^{-\tau}E_2, \tag{2.65}$$

where boundedness of E_1 and E_2 follows from boundedness of derivatives of J'. Similarly the 'f' term in (2.44) after conjugation becomes

$$G\left(e^{-\tau}f\partial_\tau\right)G^{-1} = e^{-\tau}E_3\partial_\tau + e^{-\tau}E_4. \tag{2.66}$$

Therefore the contribution to the conjugated Laplacian from terms with τ derivatives is

$$G\left(-\partial_\tau^2 - n\partial_\tau - \frac{1}{2}e^{-\tau}f\partial_\tau\right)G^{-1} = -\partial_\tau^2 + \frac{n^2}{4} + e^{-\tau}E_5\partial_\tau + e^{-\tau}E_6. \tag{2.67}$$

Now we calculate

$$G\left(e^{-2\tau}\Delta_{\gamma(0)}\right)G^{-1} = e^{-2\tau}\Delta_{\gamma(0)} + e^{-3\tau}(1+e^{-\tau}J)\left(2(\nabla_{\gamma(0)}J').\nabla_{\gamma(0)} + \Delta_{\gamma(0)}(J')\right) \tag{2.68}$$

$$= e^{-2\tau}\Delta_{\gamma(0)} + e^{-3\tau}\Omega_1.\nabla_{\gamma(0)} + e^{-3\tau}E_7,$$

and

$$G\left(e^{-3\tau}a.\nabla_{\gamma(0)}\right)G^{-1} = e^{-3\tau}a.\nabla_{\gamma(0)} + e^{-4\tau}(1+e^{-\tau}J)a.(\nabla_{\gamma(0)}J') \tag{2.69}$$

$$= e^{-3\tau}\Omega_2.\nabla_{\gamma(0)} + e^{-4\tau}E_8.$$

The final term becomes after conjugation

$$G\left(e^{-3\tau}\mathrm{div}_{\gamma(0)}.b.\nabla_{\gamma(0)}\right)G^{-1} = e^{-3\tau}\mathrm{div}_{\gamma(0)}.b.\nabla_{\gamma(0)} + e^{-4\tau}(1+e^{-\tau}J)(\nabla_{\gamma(0)}J').b.\nabla_{\gamma(0)} \tag{2.70}$$

$$\begin{aligned}
&+ e^{-4\tau}(1+e^{-\tau}J)(b.\nabla_{\gamma(0)}J').\nabla_{\gamma(0)} \\
&+ e^{-4\tau}(1+e^{-\tau}J)\mathrm{div}_{\gamma(0)}.b.(\nabla_{\gamma(0)}J') \\
&= e^{-3\tau}\mathrm{div}_{\gamma(0)}.b.\nabla_{\gamma(0)} + e^{-4\tau}\Omega_3.\nabla_{\gamma(0)} + e^{-4\tau}E_9.
\end{aligned}$$

In total therefore we have

$$\begin{aligned}
L((1),R) = G\Delta_X G^{-1} = -\partial_\tau^2 + \frac{n^2}{4} + e^{-2\tau}\Delta_{\gamma(0)} \\
+ e^{-\tau}A\partial_\tau + e^{-3\tau}B.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} \\
+ e^{-3\tau}\mathrm{div}_{\pi_{\mathcal{I}}^*\gamma(0)}.C.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} + e^{-\tau}D
\end{aligned} \tag{2.71}$$

for smooth bounded A, B, C and D. The result at a general level $\mathcal{I}$ holds with A replaced with $A_{\mathcal{I}} \equiv \pi_{\mathcal{I}}^*(A)$ and similarly for B, C and D: this follows by repeating the calculation and noting that all the inputs are lifts from level (1). The output error terms are then also lifts and the $\mathcal{I}$-uniform bounds follow. □

The following Lemma claims that the errors at each level $\mathcal{I}$ can be treated as perturbations *simultaneously*, by decreasing the size of the neighbourhood at infinity if necessary. A similar Lemma appears in [P-87, Lemma 4.1].

2.2.6 Lemma *For all $\epsilon > 0$, we can choose $R_0 = R_0(\epsilon)$ large enough so that for all $R > R_0$, and all $f \in C_0^\infty(\pi_{\mathcal{I}}^{-1}((\log R, \infty)_\tau \times B_0))$,*

$$|\langle E(\mathcal{I};R)f, f\rangle| < \epsilon|\langle L_0(\mathcal{I};R)f, f\rangle|. \tag{2.72}$$

Proof The proof will follow from the identity on $C_0^\infty(\pi_{\mathcal{I}}^{-1}((\log R,\infty)_\tau \times B_0))$

$$\langle L_0(\mathcal{I};R)f,f\rangle = \|\partial_\tau f\|^2 + \|e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f\|^2 + n^2/4\|f\|^2. \tag{2.73}$$

From Lemma 2.2.5, $E(\mathcal{I};R)$ can be written

$$E(\mathcal{I};R) = e^{-\tau}A_{\mathcal{I}}\partial_\tau + e^{-3\tau}B_{\mathcal{I}}.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} + e^{-3\tau}\mathrm{div}_{\pi_{\mathcal{I}}^*\gamma(0)}.C_{\mathcal{I}}.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)} + e^{-\tau}D_{\mathcal{I}},$$

with $A_{\mathcal{I}}, D_{\mathcal{I}}$, $\mathcal{I}$-uniformly bounded functions, $B_{\mathcal{I}}$ a vector field with $\mathcal{I}$-uniform bound and $C_{\mathcal{I}}$ representing an endomorphism of the tangent bundle. The norm of $C_{\mathcal{I}}$ is $\mathcal{I}$-uniformly bounded when considered as a function. Each of the terms occurring in $\langle E(\mathcal{I};R)f,f\rangle$ can be controlled by (2.73), e.g.

$$\begin{aligned}
|\langle e^{-3\tau}\mathrm{div}_{\pi_{\mathcal{I}}^*\gamma(0)}.C_{\mathcal{I}}.\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f,f\rangle| &= |\langle (e^{-\tau}C_{\mathcal{I}}).e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f, e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f\rangle| \\
&\leq \|(e^{-\tau}C_{\mathcal{I}}).e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f\|\|e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f\| \\
&\leq \|e^{-\tau}C_{\mathcal{I}}\|_{op}\,\|e^{-\tau}\nabla_{\pi_{\mathcal{I}}^*\gamma(0)}f\|^2 \\
&\leq \|e^{-\tau}C_{\mathcal{I}}\|_{op}\langle L_0(\mathcal{I};R)f,f\rangle \\
&\leq \frac{1}{R_0}\|C_{\mathcal{I}}\|_{op}\langle L_0(\mathcal{I};R)f,f\rangle,
\end{aligned}$$

and $\|C_{\mathcal{I}}\|_{op}$ is uniformly bounded so by increasing R_0 we can sufficiently control the size of this term. The other terms in $\langle E(\mathcal{I};R)f,f\rangle$ are estimated similarly. □

2.2.7 Lemma *For any $\eta > 0$ we can choose $R_0 = R_0(\eta)$ large enough so that for all $R > R_0$*

$$\Delta_{X(\mathcal{I})}|_{C_0^\infty(\pi_{\mathcal{I}}^{-1}((0,1/R)\times B_0))} > n^2/4 - \eta \tag{2.74}$$

uniformly through $\mathcal{I}$.

Proof By Lemma 2.2.6 we can choose $R_0(\eta)$ so that

$$|\langle E(\mathcal{I};R)f,f\rangle| < \frac{4\eta}{n^2}|\langle L_0(\mathcal{I};R)f,f\rangle|. \tag{2.75}$$

The first two terms in (2.58) are nonnegative and so we have $L_0(\mathcal{I};R) \geq n^2/4$. Then

$$\begin{aligned}\langle L(\mathcal{I};R)f,f\rangle &= \langle L_0(\mathcal{I};R)f,f\rangle + \langle E(\mathcal{I};R)f,f\rangle \\ &> (1-4\eta/n^2)\langle L_0(\mathcal{I};R)f,f\rangle \\ &> (n^2/4-\eta)\langle f,f\rangle,\end{aligned}$$

which gives the result. □

2.2.4 Eigenfunction estimates

Let $\chi : (0,\infty) \to [0,1]$ be a smooth cutoff function such that

$$\chi(t) = \begin{cases} 1 & \text{if } t \leq 1/2, \\ 0 & \text{if } t \geq 1, \end{cases}$$

and such that $(1-\chi^2)^{1/2}$ is also smooth. Then let χ_k, $k = 1,\ldots,n$ be a set of cutoff functions defined locally on $\delta\bar{X}$ such that in the boundary coordinates corresponding to $\bar{M}_k$

$$\begin{aligned}\chi_k(u,z) &= \chi(\frac{3}{2}|u|^2), \quad k < n \\ \chi_n(z) &\equiv 1\end{aligned}$$

and extended by zero to the rest of the boundary $\delta\bar{X}$. Also define scaled versions of the cutoff

$$\chi_R(\rho) = \chi(R\rho) \tag{2.76}$$

which localizes to smaller regions as $R \to \infty$. For $R > R_0$ large enough view χ_R as a function of ρ on the collar neighbourhood of infinity $(0, \epsilon)_\rho \times \delta\bar{X}$, and extend by zero to a function on the whole of X. Define for notational convenience

$$\chi_0 \equiv (1 - \sum_{k=1}^{n} \chi_k^2)^{1/2}, \tag{2.77}$$

$$\chi_{R,i,\infty} \equiv \chi_R \chi_i, \quad i = 0, 1, \ldots, n, \tag{2.78}$$

$$\chi_{R,K} \equiv (1 - \chi_R^2)^{1/2}. \tag{2.79}$$

Then the functions

$$\chi_{R,K}, \chi_{R,i,\infty}, \quad i = 0, \ldots, n$$

form an R parameterized partition of unity for X in the sense of [CFKS-87, Definition 3.1]. In particular they are appropriate for application of the IMS localization formula which first appeared explicitly in [Si-82]. Moreover the functions

$$\pi_{\mathcal{I}}^* \chi_{R,K}, \pi_{\mathcal{I}}^* \chi_{R,i,\infty}, \quad i = 0, \ldots, n$$

form an R parameterized partition of unity for $X(\mathcal{I})$. Let $K(R) = X - \rho^{-1}(0, 1/2R)$. This is a compact core for X as it is isolated from the boundary.

The bounding below in Lemmas 2.2.3 and 2.2.7 relied on the localized functions being compactly supported. In order for this to be the case we use an approximation argument following from the well known fact that on a complete Riemannian manifold (M, g), the smooth compactly supported functions $C_0^\infty(M)$ are a core for the Laplacian Δ_g. In other words $C_0^\infty(M)$ is dense in $L^2(M)$ with respect to the graph norm

$$\|f\|_{\Gamma(\Delta_g)} \equiv \sqrt{\|f\|_{L^2(M)}^2 + \|\Delta_g f\|_{L^2(M)}^2}. \tag{2.80}$$

This result can be found in the paper of Chernoff [Ch-73]. Now we state and prove the key Lemma.

2.2.8 Lemma (Eigenfunction estimates) *For any $\epsilon > 0$, there exists an $R = R(\epsilon)$ and a constant $C = C(\epsilon) > 0$ such that if ϕ is a normalized eigenfunction of the Laplacian on $X(\mathcal{I})$ with eigenvalue $s(n-s) \in [\delta(n-\delta), n^2/4 - \epsilon]$ then*

$$\int_{\pi_{\mathcal{I}}^{-1}(K(R))} |\phi|^2 dX(\mathcal{I}) \geq C > 0, \tag{2.81}$$

uniformly through $\mathcal{I}$.

Proof Suppose that ϕ is a normalized eigenfunction of $\Delta_{\mathcal{I}}$ on $X(\mathcal{I})$ with exceptional eigenvalue $s(n-s) \leq n^2/4 - \epsilon$, i.e.

$$\Delta_{\mathcal{I}}\phi = s(n-s)\phi, \quad \|\phi\|_{L^2(X(\mathcal{I}))} = 1.$$

The IMS localization formula (see [CFKS-87, Theorem 3.2] or [Si-82]) tells us how to relate global quantities to local quantities. Using this with the partition of unity $\pi_{\mathcal{I}}^*\chi_{R,K}$, $\pi_{\mathcal{I}}^*\chi_{R,i,\infty}$ for the quantity $\langle \Delta_{\mathcal{I}}\phi, \phi\rangle = s(n-s)$ we have

$$\begin{aligned} s(n-s) = \langle \Delta_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,K})\phi, (\pi_{\mathcal{I}}^*\chi_{R,K})\phi\rangle + \sum_{i=0}^{n} \langle \Delta_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\phi, (\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\phi\rangle \\ - \langle |\nabla_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,K})|^2\phi, \phi\rangle - \sum_{i=0}^{n} \langle |\nabla_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})|^2\phi, \phi\rangle. \end{aligned} \tag{2.82}$$

The first term is estimated

$$\langle \Delta_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,K})\phi, (\pi_{\mathcal{I}}^*\chi_{R,K})\phi\rangle \;\geq\; \delta(n-\delta)\|(\pi_{\mathcal{I}}^*\chi_{R,K})\phi\|^2_{L^2(X(\mathcal{I}))}, \tag{2.83}$$

by the Patterson-Sullivan description of the bottom of the spectrum. Let $\{\varphi_k\}_{k=1}^{\infty}$ be a sequence in $C_0^{\infty}(X(\mathcal{I}))$ which goes to ϕ in the graph norm. By Lemmas 2.2.3 and

2.2.7 we can increase R independently of k and $\mathcal{I}$ so that for all i

$$\langle \Delta_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\varphi_k, (\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\varphi_k\rangle \;\geq\; (n^2/4-\epsilon/4)\|(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\varphi_k\|^2_{L^2(X(\mathcal{I}))}.$$

Taking the limit in k to get the corresponding statement for ϕ and summing over i we have

$$\sum_{i=0}^{n}\langle \Delta_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\phi, (\pi_{\mathcal{I}}^*\chi_{R,i,\infty})\phi\rangle \;\geq\; (n^2/4-\epsilon/4)\|(\pi_{\mathcal{I}}^*\chi_R)\phi\|^2_{L^2(X(\mathcal{I}))}. \tag{2.84}$$

The remaining terms in (2.82) can be estimated by noting

$$\nabla_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty}) = \pi_{\mathcal{I}}^*(\nabla_{(1)}\chi_{R,i,\infty}),$$

so using the product rule for the gradient, and that the projection is a local isometry,

$$\begin{aligned}|\nabla_{\mathcal{I}}(\pi_{\mathcal{I}}^*\chi_{R,i,\infty})|^2 &= \pi_{\mathcal{I}}^*\left(\chi_R^2|\nabla_{(1)}\chi_i|^2 + \chi_i^2|\nabla_{(1)}\chi_R|^2 + 2\chi_R\chi_i\langle\nabla_{(1)}\chi_i, \nabla_{(1)}\chi_R\rangle\right)\\ &= \pi_{\mathcal{I}}^*\left(\chi_R^2|\nabla_{(1)}\chi_i|^2 + \chi_i^2|\nabla_{(1)}\chi_R|^2\right),\end{aligned}$$

where the last term on the first line vanished due to the form of the metric. Summing over i, and using the estimates (2.83) and (2.84) in (2.82) we have

$$\begin{aligned}s(n-s) \;\geq\; &\delta(n-\delta)\|(\pi_{\mathcal{I}}^*\chi_{R,K})\phi\|^2_{L^2(X(\mathcal{I}))} + (n^2/4-\epsilon/4)\|(\pi_{\mathcal{I}}^*\chi_R)\phi\|^2_{L^2(X(\mathcal{I}))} \\ &- \langle\pi_{\mathcal{I}}^*(|\nabla_{(1)}\chi_R|^2 + |\nabla_{(1)}\chi_{R,K}|^2)\phi,\phi\rangle - \langle\pi_{\mathcal{I}}^*(\chi_R^2\sum_{i=0}^{n}|\nabla_{(1)}\chi_i|^2)\phi,\phi\rangle.\end{aligned} \tag{2.85}$$

The terms on the second line will be estimated by L^∞ norms which are preserved under $\pi_{\mathcal{I}}^*$. Now we observe that $\chi_R^2|\nabla_{(1)}\chi_i|^2$ is supported only for $\rho \leq 1/R$, and there

$$\chi_R^2|\nabla_{(1)}\chi_i|_g^2 \leq |\nabla_{(1)}\chi_i|_g^2 = \rho^2|\nabla_{u,euc}\chi_i|^2_{euc} \leq 1/R^2|\nabla_{u,euc}\chi_i|^2_{euc},$$

where $\nabla_{u,euc}$, $|.|_{euc}$ refer to the Euclidean gradient and metric for the u coordinate in the regions $\bar{M}_k$. Therefore we can increase R so that

$$\|\chi_R^2 \sum_{i=0}^{n} |\nabla_{(1)}\chi_i|^2\|_\infty < \epsilon/4.$$

Incorporating these estimates we have, letting

$$F_R = (s(n-s) - \delta(n-\delta))\chi_{R,K}^2 + |\nabla_{(1)}\chi_R|^2 + |\nabla_{(1)}\chi_{R,K}|^2,$$

then

$$\begin{aligned}\|F_R^{1/2}\|_\infty^2 \|\phi\|^2_{|L^2(\pi_{\mathcal{I}}^{-1}(K(R)))} &\geq \|(\pi_{\mathcal{I}}^* F_R)^{1/2}\phi\|^2_{L^2(X(\mathcal{I}))} \\ &\geq (n^2/4 - \epsilon/2 - s(n-s))\|\phi\|^2_{L^2(X(\mathcal{I})-\pi_{\mathcal{I}}^{-1}(K(R)))}.\end{aligned}$$

Now we note that $\|F_R^{1/2}\|_\infty$ is uniformly bounded as $R \to \infty$. This follows from $s(n-s) < n^2/4$, and for example

$$|\nabla_{(1)}\chi_R|_g^2 = |\rho^2 R\chi'|_g^2 = \rho^2 R^2 |\chi'|^2 \leq |\chi'|^2,$$

where $'$ denotes a derivative. Then we have shown

$$\begin{aligned}1 &= \|\phi\|^2_{|L^2(\pi_{\mathcal{I}}^{-1}(K(R)))} + \|\phi\|^2_{L^2(X(\mathcal{I})-\pi_{\mathcal{I}}^{-1}(K(R)))} \\ &\leq \left(1 + \frac{\|F_R^{1/2}\|_\infty^2}{(n^2/4 - \epsilon/2 - s(n-s))}\right) \|\phi\|^2_{|L^2(\pi_{\mathcal{I}}^{-1}(K(R)))} \\ &\leq \left(1 + \frac{\|F_R^{1/2}\|_\infty^2}{\epsilon/2}\right) \|\phi\|^2_{|L^2(\pi_{\mathcal{I}}^{-1}(K(R)))} \\ &\leq (1 + C_0(\epsilon)) \|\phi\|^2_{|L^2(\pi_{\mathcal{I}}^{-1}(K(R)))} \\ &= (1 + C_0(\epsilon)) \int_{\pi_{\mathcal{I}}^{-1}(K(R))} |\phi|^2 dX(\mathcal{I})\end{aligned}$$

which establishes the result by taking

$$C(\epsilon) = 1/(1 + C_0(\epsilon)) > 0.$$

□

2.3 Analytic Preparations

2.3.1 Estimates for terms appearing in the trace formula

Let o denote the point corresponding to $(0, \ldots, 0, 1)$ in the hyperboloid model for $\mathbb{H}^{n+1}$. We write $G = \mathrm{SO}^0(n+1, 1)$, $K = \mathrm{SO}(n+1)$ the maximal compact subgroup. T will be some generating parameter throughout the rest of the paper. For the details on spherical functions the reader can see [H-08]. For the representation theory we refer to [Kn-86]. All further ideas in this section are due to Sarnak and Xue [SX-91].

The lattice point count relates to harmonic analysis by consideration of the function $\chi_T : G \to \mathbb{R}$

$$\chi_T(g) = \begin{cases} 1 & \text{if } d(o, g(o)) \leq T, \\ 0 & \text{if } d(o, g(o)) > T. \end{cases} \tag{2.86}$$

For any $\lambda = s(n-s) \in (0, n^2/4)$ we can consider the associated compliments series representation π_s. This contains a normalized spherical (K-invariant) vector v. From this data we construct the spherical function

$$\phi_s(g) \equiv \langle \pi_s(g)v, v \rangle. \tag{2.87}$$

We let $f_s(g) = \chi_T(g)\phi_s(g)$, and $F_s = f_s * \overline{f_s}$ with $\overline{f_s}(g) = \overline{f_s(g^{-1})}$. Then following

[SX-91, Lemma 2.1] we have $F_s \in C_0(K\backslash G/K)$ and

$$F_s(g) \ll \begin{cases} e^{2(s-n/2)T} e^{-\frac{n}{2}d(o,g(o))} & \text{if } d(o,g(o)) \leq 2T, \\ 0 & \text{if } d(o,g(o)) > 2T. \end{cases} \tag{2.88}$$

The implied constant can be taken uniformly for $s \in I \subset (n/2, \delta]$, for I a closed interval. As F_s is K-biinvariant there is an associated spherical transform of F_s, defined for $\lambda_t = t(n-t)$

$$\hat{F}_s(\lambda_t) = \int_G F_s(g)\overline{\phi_t(g)}dG. \tag{2.89}$$

In fact, the spherical transform is a $*$-homomorphism so we have, evaluating at $\lambda = \lambda_s$

$$\hat{F}_s(\lambda) = |f_s(\lambda)|^2 = \left(\int_G \chi_T(g)|\phi_s(g)|^2 dG\right)^2, \tag{2.90}$$

and using the property

$$\phi_s(\exp tX) \gg e^{(s-n)t} \tag{2.91}$$

we have

$$\hat{F}_s(\lambda) = \left(\int_G \chi_T(g)|\phi_s(g)|^2 dG\right)^2 \gg \left(\int_0^T e^{2(s-n)t} e^{nt} dt\right)^2 \gg e^{4(s-n/2)T}. \tag{2.92}$$

Moreover, as before the implied constant can be chosen uniformly for s in compact $I \subset (n/2, \delta]$.

2.3.2 Lattice point count

Now we estimate the quantity

$$N(\Gamma(\mathcal{I}), T) \equiv |\{\gamma \in \Gamma(\mathcal{I}) : d(o, \gamma o) \leq T\}|. \tag{2.93}$$

In [SX-91] Sarnak and Xue conjectured that

$$N(\Gamma(\mathcal{I}), T) \ll_\epsilon \frac{e^{nT(1+\epsilon)}}{[\Gamma : \Gamma(\mathcal{I})]} + e^{nT/2} \tag{2.94}$$

for Γ an arithmetic lattice in $\mathrm{SO}(n+1,1)$. They established this result for $n=1,2$ by a direct counting argument.

For $n \geq 3$ we will rely on a result of Kelmer and Silberman from [KeSi-10]. This uses the spectral theory at the cofinite level (for Γ). The best known spectral gap when $n \geq 3$ is given by Theorem 1.0.3 (a result of Bergeron and Clozel from [BC-12]). This tells us that if $s > n/2$ and $s(n-s)$ is a nonzero eigenvalue for $\mathbb{H}^{n+1}/\Gamma(\mathcal{I})$ then $s \leq n-1$.

The consequence for the lattice point count in $\mathrm{SO}(n+1,1)$, $n \geq 3$ is

$$N(\Gamma(\mathcal{I}),T) \ll \frac{e^{nT}}{[\Gamma:\Gamma(\mathcal{I})]} + e^{(n-1)T} \tag{2.95}$$

uniformly in $T, \mathcal{I}$. This result appears in [KeSi-10, Theorem 2]. Using the estimate on the size of the factor group from Lemma 2.1.5, we have the following bound for the lattice point count.

2.3.1 Lemma (Lattice point count) *For any $\epsilon > 0$ and $n \geq 2$ we have*

$$N(\Gamma(\mathcal{I}),T) \ll_\epsilon \frac{e^{nT(1+\epsilon)}}{|\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}} + e^{(n-1)T}. \tag{2.96}$$

2.4 Proof of Main Theorem

Let $I = [a,\delta]$ a closed interval for some $a > n/2$. Replace Λ with the Λ_1 of Section 2.1 if necessary. We aim to apply the pre-trace formula to the automorphic kernel on $\mathbb{H}^{n+1} \times \mathbb{H}^{n+1}$ corresponding to F_s at level $\mathcal{I}$ i.e.

$$K_{\mathcal{I}}(x_1,x_2) = \sum_{\gamma \in \Lambda(\mathcal{I})} F_s(g_{x_1}^{-1}\gamma g_{x_2}), \tag{2.97}$$

where we write g_x for any group element such that $g_x(o) = x$. We have for the spectral decomposition of the automorphic kernel, as in [Ga-02, Proposition 5.2],

$$K_{\mathcal{I}}(x,x) = \sum_{\lambda_{j,\mathcal{I}} < n^2/4} \hat{F}_s(\lambda_{j,\mathcal{I}}) |\psi_j(x)|^2 + \mathcal{E}, \tag{2.98}$$

where $\mathcal{E}$ is some *nonnegative* contribution from the continuous spectrum. The $\lambda_{j,\mathcal{I}}$ are the eigenvalues of the Laplacian on $X(\mathcal{I})$ below $n^2/4$, counted with multiplicities. The ψ_i are the corresponding (lifted) eigenfunctions. There are only finitely many such eigenfunctions by the work of Lax and Phillips [LP-82].

We can now apply the eigenfunction estimates (Lemma 2.2.8) to find a compact part $\mathcal{K} \subset X$ such that for all $s_i \in I$

$$\int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} |\psi_j(x)|^2 dX(\mathcal{I}) \geq C > 0 \tag{2.99}$$

uniformly through $s_i \in I$ and $\mathcal{I}$. This implies

$$\begin{aligned} \int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} K_{\mathcal{I}}(x,x) dX(\mathcal{I}) &\geq \sum_{\lambda_{j,\mathcal{I}} < n^2/4} \hat{F}_s(\lambda_{j,\mathcal{I}}) \int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} |\psi_j(x)|^2 \\ &\geq \sum_{\lambda_{j,\mathcal{I}} : s_j \in I} \hat{F}_s(\lambda_{j,\mathcal{I}}) \int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} |\psi_j(x)|^2 \\ &\geq C \sum_{\lambda_{j,\mathcal{I}} : s_j \in I} \hat{F}_s(\lambda_{j,\mathcal{I}}). \end{aligned}$$

In particular, for any $s \in I$, if $\lambda = s(n-s)$ appears as an eigenvalue of $\Delta_{X(\mathcal{I})}$

$$\int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} K_{\mathcal{I}}(x,x) dX(\mathcal{I}) \gg \hat{F}_s(\lambda) \gg e^{4(s-n/2)T} m(\lambda, \mathcal{I}), \tag{2.100}$$

where $m(\lambda, \mathcal{I})$ is the multiplicity of the eigenvalue and the implied constant is uniform through all $s \in I$. The last inequality is a result of the estimate for the spherically transformed kernel in (2.92).

On the other hand

$$\begin{aligned}\int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} K_{\mathcal{I}}(x,x)dX(\mathcal{I}) &= \sum_{\gamma\in\Lambda(\mathcal{I})}\int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} F_s(g_x^{-1}\gamma g_x)dX(\mathcal{I}) \\ &= \sum_{\gamma\in\Lambda(\mathcal{I})}\sum_{l\in\Lambda/\Lambda(\mathcal{I})}\int_{\mathcal{K}} F_s(g_x^{-1}l^{-1}\gamma l g_x)dX \\ &\ll |\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}\sum_{\gamma\in\Lambda(\mathcal{I})}\int_{\mathcal{K}} F_s(g_x^{-1}\gamma g_x)dX \\ &\ll |\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}\sum_{\gamma\in\Gamma(\mathcal{I})}\int_{\mathcal{K}} F_s(g_x^{-1}\gamma g_x)dX.\end{aligned}$$

The penultimate inequality uses the bound on the size of the group $\Lambda/\Lambda(\mathcal{I})$ given in Lemma 2.1.5. The last inequality is a result of the rather crude observation that $\Lambda(\mathcal{I}) \subset \Gamma(\mathcal{I})$. Using the upper bound (2.88) we have

$$\sum_{\gamma\in\Gamma(\mathcal{I})}\int_{\mathcal{K}} F_s(g_x^{-1}\gamma g_x)dX \ll e^{2(s-n/2)T}\sum_{\gamma\in\Gamma(\mathcal{I})}\int_{x\in\mathcal{K}:d(x,\gamma x)\leq 2T} e^{-\frac{n}{2}d(x,\gamma x)}dX. \quad (2.101)$$

As $\mathcal{K}$ is compact there exists R such that $\mathcal{K} \subset B(o,R)$ in $\mathbb{H}^{n+1}$. This gives

$$d(o,\gamma o) \leq d(o,x) + d(x,\gamma x) + d(\gamma x,\gamma o) \leq 2R + d(x,\gamma x),$$

and

$$d(x,\gamma x) \leq 2R + d(o,\gamma o).$$

Then

$$\begin{aligned}\sum_{\gamma\in\Gamma(\mathcal{I})}\int_{x\in\mathcal{K}:d(x,\gamma x)\leq 2T} e^{-\frac{n}{2}d(x,\gamma x)}dX &\ll \sum_{\gamma\in\Gamma(\mathcal{I}):d(o,\gamma o)\leq 2T+2R} e^{-\frac{n}{2}d(o,\gamma o)}dX \\ &\ll \int_0^{2T+2R} e^{-\frac{n}{2}t}N(\Gamma(\mathcal{I}),t)dt,\end{aligned}$$

by integrating by parts. By using the lattice point bound from Lemma 2.3.1 we

estimate

$$\int_0^{2T+2R} e^{-\frac{n}{2}t} N(\Gamma(\mathcal{I}),t)dt \ll_\epsilon \int_0^{2T+2R} \frac{e^{nt(1/2+\epsilon)}}{|\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}} + e^{(n/2-1)t}$$
$$\ll_\epsilon \frac{e^{nT(1+2\epsilon)}}{|\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}} + e^{(n-2)T}.$$

Gathering together we have the upper bound

$$\int_{\pi_{\mathcal{I}}^{-1}(\mathcal{K})} K_{\mathcal{I}}(x,x)dX(\mathcal{I}) \ll_\epsilon e^{2(s-n/2)T}\left(e^{nT(1+2\epsilon)} + |\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}e^{(n-2)T}\right), \tag{2.102}$$

so that we now have upper and lower bounds for the partial trace. Equations (2.102) and (2.100) give

$$e^{2(s-n/2)T} m(\lambda,\mathcal{I}) \ll_\epsilon e^{nT(1+2\epsilon)} + |\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}e^{(n-2)T}.$$

This is the keystone of the proof. Our previous work on the multiplicities and the lattice point count will now play together to forbid certain values of s. Using Lemma 2.1.5 to estimate $m(\lambda,\mathcal{I})$ we now have for any $\epsilon > 0$

$$e^{2(s-n/2)T}|\mathcal{O}_F/\mathcal{I}|^{n-1} \ll_\epsilon e^{nT(1+2\epsilon)} + |\mathcal{O}_F/\mathcal{I}|^{(n+2)(n+1)/2}e^{(n-2)T}. \tag{2.103}$$

Taking $T \approx ((n+1)(n+2)/4)\log|\mathcal{O}_F/\mathcal{I}|$ we have for all $\epsilon > 0$

$$|\mathcal{O}_F/\mathcal{I}|^{(2s-n)(n+1)(n+2)/4+n-1} \ll_\epsilon |\mathcal{O}_F/\mathcal{I}|^{(1+2\epsilon)n(n+1)(n+2)/4} \tag{2.104}$$

as $|\mathcal{O}_F/\mathcal{I}| \to \infty$, which can only be true if

$$s \leq s_n^0 \equiv n - \frac{2(n-1)}{(n+1)(n+2)}.$$

Recall that at the start we were free to choose an interval I_s throughout which we had uniformity. Going back and choosing

$$I = [s_n^0, \delta],$$

which makes sense as long as $\delta > s_n^0$, we have proved Theorem 1.0.6. Corollary 1.0.7 follows directly.

Chapter 3

Construction of thin groups with thick limit sets

Our main theorem (Theorem 1.0.6) gives a quantitative spectral gap for infinite covolume subgroups of hyperbolic isometries, provided the Hausdorff dimension of the limit set is large enough (and some other conditions are met). In other words, the 'bass note' of the quotient manifold must be low. In order to find such a group, we must remain safe from Doyle's pretty result [D-88], which states that there is a universal upper bound for the Hausdorff dimension of the limit sets of Schottky groups in $\mathrm{Isom}(\mathbb{H}^3)$.

The following procedure for constructing thin subgroups of hyperbolic arithmetic lattices with arbitrarily large Hausdorff dimension is due to McMullen [McM-12]. He has communicated a method which takes as input a compact $n+1$ dimensional hyperbolic manifold M which arises as an arithmetic quotient $\mathbb{H}^{n+1}/\Gamma$ and which has an

embedded totally geodesic hypersurface S such that $[S] \neq 0$ in $H_n(M, \mathbf{Z})$. The output is a set of subgroups of Γ which are geometrically finite, infinite index and have Hausdorff dimensions of the limit set arbitrarily close to n. This yields groups for which our main theorem (Theorem 1.0.6) applies. This construction is similar in nature to the construction of Gamburd in the last section of [Ga-02]. Slightly more care is required in higher dimension, for example because of the inequivalence of geometrical finiteness and finite generation.

The required input arises in the work of Millson [Mill-76], whose development we recount now. We take as our field $F = \mathbf{Q}(\sqrt{p})$ with p prime and consider the quadratic form

$$q(x_1, \ldots, x_{n+2}) = x_1^2 + x_2^2 + \ldots x_{n+1}^2 - \sqrt{p}\, x_{n+2}^2. \tag{3.1}$$

We then take Γ_0 to be the subgroup of $\mathrm{GL}_{n+2}(\mathcal{O}_F)$ which preserves q. The quadratic form q is conjugate to $\mathrm{diag}(1, \ldots, 1, -1)$ by a diagonal matrix over $\mathbb{R}$, via conjugation with the same element we can realize Γ_0 as a group of isometries of the hyperbolic plane $\mathbb{H}^{n+1}$. We let Γ be the orientation preserving subgroup of Γ_0. As we have remarked in a previous section, Γ is discrete. Moreover Γ is uniform (cocompact) as follows. There are no rational zeros of q, and Γ contains no nontrivial unipotent elements (if $\gamma \in \Gamma$ is unipotent then so is its Galois conjugate, which is an element of a definite orthogonal group). The work of Mostow and Tamagawa [MT-62] then gives uniformity of Γ. It is clear that Γ is geometrically finite.

We next consider the congruence subgroups $\Gamma(\mathcal{P})$ for $\mathcal{P}$ prime in F. As in [Mill-76] for $\mathcal{P}$ of large enough norm $\Gamma(\mathcal{P})$ is torsion-free and the quotient $\mathbb{H}^{n+1}/\Gamma(\mathcal{P})$ is

a compact manifold of constant negative curvature. Millson considers the involution ι in Γ which is the reflection in the plane $x_1 = 0$. As $\Gamma(\mathcal{P})$ is normal in Γ then ι normalizes $\Gamma(\mathcal{P})$ and hence descends to an involution on $\mathbb{H}^{n+1}/\Gamma(\mathcal{P})$. For $\mathcal{P}$ not of norm two and such that $\Gamma(\mathcal{P})$ is torsion free we have then a manifold $M(\mathcal{P}) = \mathbb{H}^{n+1}/\Gamma(\mathcal{P})$ with an involution ι whose fixed point set contains a constantly negatively curved, orientable, codimension one submanifold $S(\mathcal{P})$. Millson goes on to show that by passing to congruence subgroups

$$\Gamma(\mathcal{P}, \mathcal{P}') \equiv \{\gamma \in \Gamma : \gamma \equiv I \mod \mathcal{P},\ \gamma \equiv I \mod \mathcal{P}'\} \tag{3.2}$$

with $\mathcal{P}$ and $\mathcal{P}'$ of large enough norm we obtain a manifold $M = \mathbb{H}^{n+1}/\Gamma(\mathcal{P}, \mathcal{P}')$ and a corresponding embedded totally geodesic hypersurface S such that $[S] \neq 0 \in H_n(M, \mathbf{Q})$. We have $\pi_1(M) = \Gamma(\mathcal{P}, \mathcal{P}')$ a geometrically finite and finite index subgroup of a uniform arithmetic lattice.

We are now ready to apply McMullen's construction [McM-12].

Let $\pitchfork : H_k(M, \mathbf{Z}) \times H_{n+1-k}(M, \mathbf{Z}) \to H_0(M, Z) \cong \mathbf{Z}$ denote intersection number and consider the homomorphism

$$\phi : \pi_1(M) \to \mathbf{Z}, \quad \phi(\gamma) = [S] \pitchfork [\gamma], \tag{3.3}$$

by using the usual map $\pi_1(M) \to H_1(M, \mathbf{Z})$. Poincaré duality gives that ϕ is defined and nonzero (the reader can see Hatcher [Ha-02] for details), as S is oriented and embedded ϕ is moreover onto. The kernel $\ker \phi$ consists of classes with generic representatives which cross S the same number of times in each direction relative to a fixed orientation of S.

We can cut M open along S to give a connected manifold with boundary M_{cut}. The boundary δM_{cut} consists of two connected components S_1 and S_2 which are both isometric to S. Recall that S was obtained as a quotient of a totally geodesic hyperplane $\Pi \subset \mathbb{H}^{n+1}$. In this case it is known that $\pi_1(M) = \pi_1(M_{cut}) *_{\pi_1(S)}$, the HNN extension of $\pi_1(M_{cut})$ with respect to the maps $i_{j*} : \pi_1(S_j) \to \pi_1(M_{cut})$ induced from the inclusions of the boundary components of M_{cut}. Formally the HNN extension is generated by $\pi_1(M_{cut})$ and an element t subject to the relations

$$t i_1(\gamma) t^{-1} = i_2(\gamma), \tag{3.4}$$

concretely this element t can be chosen to be anything with $\phi(t) = 1$. Fix such a t.

To obtain manifolds with many repeated sections we consider the collection of hyperplanes

$$\mathcal{C}_k = \{\gamma\Pi : \gamma \in \phi^{-1}(0) = \ker \phi\} \cup \{\gamma\Pi : \gamma \in \phi^{-1}(k)\}. \tag{3.5}$$

Distinct hyperplanes in this collection do not intersect, and there is a connected component of $\mathbb{H}^{n+1} - \mathcal{C}_k$ which meets all of the hyperplanes in $\mathcal{C}_k$, we write R_k for the closure of this connected component. We let $\Gamma_k = \mathrm{Stab}_{\pi_1(M)}(R_k)$ so that for $k = 1$ we have that $\Gamma_1 = \pi_1(M_{cut})$ and R_1 a universal cover for M_{cut}, i.e. $M_{cut} = R_1/\Gamma_1$.

The quotient R_k/Γ_k is then the manifold with boundary which is obtained by gluing copies M^i_{cut}, $i = 1, \ldots, k$ together by identifying one boundary component of M^i_{cut} with the opposite boundary component of M^{i+1}_{cut}, leaving two boundary components unglued (coming from the first and last copies of M_{cut}).

We now note some properties of Γ_k. As Γ_k is contained in the kernel of ϕ it is of infinite index in $\pi_1(M)$. As R_k is invariant under Γ_k the limit set $L(\Gamma_k)$ of

Γ_k must lie between the boundary of Π and the boundary of $t^k\Pi$. This implies that the quotient $CL(\Gamma_k)/\Gamma_k$ of the closed convex hull of $L(\Gamma_k)$ by Γ_k has finite volume (in fact it is compact). Also noting that Γ_k is finitely generated implies that Γ_k is geometrically finite by a result of Bowditch [Bow-93].

We have constructed Γ_k, an infinite index geometrically finite subgroup of an arithmetic lattice in $\mathrm{SO}(n+1,1)$. It remains to show that by making k large enough we can force the Hausdorff dimension of the limit set $\delta(L(\Gamma_k))$ to be as close to n as we like. Equivalently by the work of Sullivan [Sul-82] we can show that there are arbitrarily small eigenvalues of the Laplacian on $X_k \equiv \mathbb{H}^{n+1}/\Gamma_k$ for k large. This will follow by finding functions u_k with small Rayleigh quotient

$$\frac{\int_{X_k} \|\nabla u_k\|^2 dX_k}{\int_{X_k} |u_k|^2 dX_k}. \tag{3.6}$$

We have R_k/Γ_k a closed subset of X_k, consisting of k glued copies M^i_{cut} of M_{cut}. Let $f_+ \in C^\infty(M_{cut})$ be such that $f_+|_{S_1} \equiv 0$ and $f_+|_{S_2} \equiv 1$ and which is locally constant in a neighbourhood of the boundary. Now we define for $k \geq 2$

$$u_k = \begin{cases} f_+ & \text{on } M^1_{cut} \\ 1 & \text{on } M^i_{cut},\ i = 2, \ldots, k-1 \\ (1 - f_+) & \text{on } M^k_{cut} \\ 0 & \text{on } X_k - R_k/\Gamma_k. \end{cases} \tag{3.7}$$

Calculation of the Rayleigh quotient (3.6) gives

$$\frac{\int_{X_k} \|\nabla u_k\|^2 dX_k}{\int_{X_k} |u_k|^2 dX_k} = \frac{2\int_{M_{cut}} \|\nabla f_+\|^2 dM_{cut}}{\int_{M_{cut}} (f_+)^2 + (1-f_+)^2 dM_{cut} + (k-2)\mathrm{vol}(M_{cut})} \tag{3.8}$$

which tends to 0 as $k \to \infty$ (the only term depending on k is the last summand in the denominator). This completes the construction.

Bibliography

[A-05] J. Arthur. An introduction to the trace formula, in *Harmonic analysis, the trace formula, and Shimura varieties*, Clay Math. Proc., vol **4**:1-263, Amer. Math. Soc., Providence, RI, 2005.

[Bass-74] H. Bass. Clifford Algebras and Spinor Norms Over a Commutative Ring, American Journal of Mathematics **96**, no. 1:156-206, 1974.

[Bor-60] A. Borel. Density Properties for Certain Subgroups of Semi-Simple Groups Without Compact Components, Annals of Mathematics **72**:179-188, 1960.

[BB-11] V. Blomer, F. Brumley. On the Ramanujan conjecture over number fields, Preprint, arXiv:math.NT/1003.0559, 2011.

[BC-12] N. Bergeron, L. Clozel. Quelques conséquences des travaux d'Arthur pour le spectre et la topologie des variétés hyperboliques, Invent. math. (Electronic publication), DOI 10.1007/s00222-012-0415-2, 2012.

[BG-08] J. Bourgain, A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$, Annals of Mathematics **167**:625-642, 2008.

[BGS-06] J. Bourgain, A. Gamburd, P. Sarnak. Sieving and Expanders, C. R. Acad. Sci. Paris, Ser. I **343**:155-159, 2006.

[BGS-09] J. Bourgain, A. Gamburd, P. Sarnak. Generalization of Selberg's 3/16 theorem and affine sieve, Acta Math. **207**:255-290, 2011.

[BGS-10] J. Bourgain, A. Gamburd, P. Sarnak. Affine linear sieve, expanders, and sum-product, Invent. math. **179**:559-644, 2010.

[Bow-93] B.H. Bowditch. Geometrical Finiteness for Hyperbolic Groups, J. Funct. Anal. **113**:245-317, 1993.

[BS-91] M. Burger, P. Sarnak. Ramanujan Duals II, Invent. math. **106**, no. 1:1-11, 1991.

[CFKS-87] H.L. Cycon, R.G. Froese, W. Kirsch, B. Simon. *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*, Springer-Verlag, Berlin-Heidelberg-New York, 1987.

[Ch-73] P. Chernoff. Essential self-adjointness of powers of generators of hyperbolic equations, J. Funct. Anal. **12**:401-414, 1973.

[CL-10] P.-H. Chaudouard, G. Laumon. Le lemme fondamental pondéré. I. Constructions géométriques, Compos. Math. **146**, no. 6:1416-1506, 2010.

[CL-09] P.-H. Chaudouard, G. Laumon. Le lemme fondamental pondéré. II. Énoncés cohomologiques, Preprint, arXiv:math.AG/0912.4512, 2009.

[Clo-03] L. Clozel. Démonstration de la conjecture τ, Invent. math. **151**:297-328, 2003.

[CLPS-91] J. Cogdell, J.-S. Li, I. Piatetski-Shapiro, P. Sarnak. Poincaré series for $\mathrm{SO}(n,1)$, Acta Math. **167**:229-285, 1991.

[D-88] P.G. Doyle. On the bass note of a Schottky group, Acta Math. **160**:249-284, 1988.

[EGM-90] J. Elstrodt, F. Grunewald, J. Mennicke. Kloosterman sums for Clifford algebras and a lower bound for the positive eigenvalues of the Laplacian for congruence subgroups acting on hyperbolic spaces, Invent. math. **101**:641-685, 1990.

[F-62] J. M. G. Fell. Weak containment and induced representation of groups, Canadian Journal of Mathematics **14**:237-268, 1962.

[Ga-02] A. Gamburd. On the Spectral Gap for Infinite Index "Congruence" Subgroups of $\mathrm{SL}_2(\mathbf{Z})$, Israel Journal of Mathematics **127**:157-200, 2002.

[GJ-78] S. Gelbart, H. Jacquet. A relation between automorphic representations of GL(2) and GL(3), Annales Scientifiques de l'École Normale Supérieure **11**:471-552, 1978.

[Gu-09] C. Guillarmou. Scattering theory on geometrically finite quotients with rational cusps, Cubo **11**, no. 5:129-172, 2009.

[Ha-02] A. Hatcher. *Algebraic Topology*, Cambridge University Press, 2002.

[H-08] S. Helgason. *Geometric Analysis on Symmetric Spaces*, Mathematical Surveys and Monographs, Second Edition, Amer. Math. Soc., 2008.

[I-89] H. Iwaniec. Selberg's lower bound of the first eigenvalue of congruence subgroups, in *Number Theory, Trace Formulas, Discrete Groups* (E.J. Alvers, ed.), Academic Press, New York, 371-375, 1989.

[I-96] H. Iwaniec. The lowest eigenvalue for congruence groups, Progress in Nonlinear Differential Equations and Applications (Boston, Mass.); vol. **20**:203-212, Birkhäuser, 1996.

[Ka-00] M. Kapovich. *Hyperbolic Manifolds and Discrete Groups*, Progress in Mathematics (Boston, Mass.); vol **183**, Birkhäuser, 2000.

[Kn-86] A. Knapp. *Representation Theory of Semisimple Groups, An Overview Based on Examples*, Princeton University Press, 1986,

[KeSi-10] D. Kelmer, L. Silberman. A Uniform Spectral Gap for Congruence Covers of a Hyperbolic Manifold, Preprint, arXiv:math.NT/1010.1010, 2010.

[KiSa-03] H.H. Kim, P. Sarnak. Refined estimates towards the Ramanujan and Selberg conjectures. Appendix to H.H. Kim, J. Amer. Math. Soc. **16**, no. 1:139-183, 2003.

[KiSh-02] H.H. Kim, F. Shahidi. Products for $\mathrm{GL}_2 \times \mathrm{GL}_3$ and the Symmetric Cube for GL_2, Annals of Mathematics, **155**:837-893, 2002.

[Ko-09] A.V. Kontorovich. The hyperbolic lattice point count in infinite volume with applications to sieves, Duke Mathematical Journal **149**, no. 1:1-36, 2009.

[KoOh-10] A. Kontorovich, H. Oh. Almost prime Pythagorean triples in thin orbits, Preprint, arxiv:math.NT/1001.0370v1, 2010.

[LP-82] P.D. Lax, R.S. Phillips. The asymptotic distribution of lattice points in Euclidean and non-Euclidean space, Journal of Functional Analysis, **46**:280-350, 1982.

[LRS-95] W. Luo, Z. Rudnick, P. Sarnak. On Selberg's eigenvalue conjecture, Geometric and Functional Analysis **5**, no. 2:387-401, 1995.

[Lu-93] A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*, Progress in Mathematics (Boston, Mass.); vol. **195**, Birkhäuser, 1993.

[McM-12] C. McMullen. Private Communication (Email) 02/19/2012.

[Mill-76] J.J. Millson. On the First Betti Number of a Constant Negatively Curved Manifold. Annals of Mathematics, **104**, no. 2:235-247, 1976.

[Mi-65] J.W. Milnor. *Topology from the Differentiable Viewpoint*, University of Virginia Press, Charlottesville, 1965.

[MP-90] R. Mazzeo, R.S. Phillips. Hodge Theory on Hyperbolic Manifolds. Duke Mathematical Journal **60**, no. 2:509-559, 1990.

[MT-62] G.D. Mostow, T. Tamagawa. On the Compactness of Arithmetically Defined Homogeneous Spaces. Annals of Mathematics, **76**, no. 3:446-463, 1962.

[N-10] B. C. Ngô. Le lemme fondamental pour les algèbres de Lie, Publ. Math. Inst. Hautes Études Sci. **111**:1-169, 2010.

[P-87] P.A. Perry. The Laplace Operator on a Hyperbolic Manifold I. Spectral and Scattering Theory. Journal of Functional Analysis **75**:161-187, 1987.

[Sa-83] P. Sarnak. The arithmetic and geometry of some hyperbolic three manifolds, Acta Math. **151**:253-295, 1983.

[Sa-95] P. Sarnak. Selberg's eigenvalue conjecture, Notices of the American Mathematical Society **42**:1272-1277, 1995.

[Sa-05] P. Sarnak. Notes on the Generalized Ramanujan Conjectures, Clay Mathematics Proceedings, **4**:659-685, 2005.

[Se-60] A. Selberg. On Discontinuous Groups in Higher Dimensional Symmetric Spaces, Contributions to Function Theory, 147-164, Tata Institute of Fundamental Research, Bombay, 1960.

[Se-65] A. Selberg. On the estimation of Fourier coefficients of modular forms, Proceedings of Symposia in Pure Mathematics **VIII**, American Mathematical Society, 1-15, 1965.

[Si-82] I.M. Sigal. Geometric Methods in the Quantum Many-Body Problem. Nonexistence of Very Negative Ions, Commun. Math. Phys., **85**:309-324, 1982.

[Suz-82] M. Suzuki. *Group Theory I*, Springer-Verlag, Berlin-Heidelberg-New York, 1982.

[Sul-82] D. Sullivan. Discrete conformal groups and measurable dynamics, Bulletin of the American Mathematical Society **6**:57-73, 1982.

[SX-91] P. Sarnak, X.X. Xue. Bounds for multiplicities of automorphic representations, Duke Math. J. **64**, no. 1:207-227, 1991.

[SZ-93] G.M. Seitz, A.E. Zalesskii. On the Minimal Degrees of Projective Representations of the Finite Chevalley Groups, II. Journal of Algebra **158**:233-243, 1993.

[T-97] W. Thurston. *The Geometry and Topology of 3-Manifolds*, Princeton University Press, 1997.

[Weil-48] A. Weil. On some exponential sums, Proceedings of the National Academy of Sciences of the United States of America, **34**:204-207, 1948.

[Weis-84] B. Weisfeiler. Strong approximation for Zariski-dense subgroups of semi-simple algebraic groups, Annals of Mathematics, **120**:271-315, 1984.