HYBRID BOUNDS FOR AUTOMORPHIC FORMS ON ELLIPSOIDS OVER NUMBER FIELDS

VALENTIN BLOMER AND PHILIPPE MICHEL
NOVEMBER 21, 2012

Abstract. We prove upper bounds for Hecke-Laplace eigenfunctions on certain Riemannian mani-

folds \(X\) of arithmetic type, uniformly in the eigenvalue and the volume of the manifold. The

manifolds under consideration are \(d\)-fold products of 2-spheres or 3-spheres, realized as adelic

quotients of quaternion algebras over totally real number fields. In the volume aspect we prove a

(“Weyl-type”) saving of \(\text{vol}(X)^{-1/6+\varepsilon}\).

Contents

1. Introduction 1
2. Preliminaries 5
3. Reduction of definite quadratic forms 12
4. Representation numbers of quadratic forms 13
5. Application of the pre-trace formula 18
6. Bounds for automorphic forms on 3-dimensional ellipsoids 22
References 24

1. INTRODUCTION

Given a Riemannian manifold \(X\), it is a classical problem to give a pointwise upper bound for a \(L^2\)-normalized Laplace eigenfunction \(\phi\) in terms of the Laplace eigenvalue \(\lambda\) and/or properties of \(X\).

It can be either seen as a rough measure of the non-concentration of the mass of \(\phi\) or as a degenerate restriction problem (where the cycle is reduced to a single point). If \(X\) is compact, generic methods give the bound [Sar]

\[
\|\phi\|_\infty \ll_X (1 + |\lambda|)^{(\dim X - 1)/4},
\]

and one seeks improvements over this bound. As pointed out in [Sar], the sup-norm problem is also closely tied to the multiplicity problem: If \(V_\lambda\) denotes the eigenspace of the eigenvalue \(\lambda\), generic methods show that

\[
\dim V_\lambda \leq \text{vol}(X) \sup_{\|\phi\|_\infty = 1} \|\phi\|^2.
\]

In other words, high multiplicity of eigenvalues may be an obstruction to sup-norm bounds better than (1.1), at least for general eigenfunctions \(\phi\). For instance, in the case of the sphere \(X = S^2\), the dimension of the eigenvalue \(\lambda = k(k + 1) (k \in \mathbb{N}_0)\) is known to be \(2k + 1\). Hence by (1.2) the best possible sup-norm bound we can hope for is of order \((1 + k)^{1/2} \approx (1 + \lambda)^{1/4}\) which is realized by the

2000 Mathematics Subject Classification. Primary 11R52, 11F70, 11F72, 11E20, Secondary: 58J50.
Key words and phrases. definite quaternion algebras, trace formula, sup-norms, hybrid bounds, norm forms, spherical harmonics.

V. B. is supported by a Volkswagen Lichtenberg fellowship and an ERC starting grant. Ph. M. is partially supported by the ERC advanced research grant n. 228304 and the SNF grant 200021-12529.
\(L^2\)-normalized zonal spherical function \(\sqrt{(2k+1)/(4\pi)}p_k(\cos \theta)\) (where \(\theta\) is the polar angle and \(p_k\) is the \(k\)-th Legendre polynomial). This situation is in sharp contrast to negatively curved Riemann surfaces where the sup-norms of eigenfunctions (and in particular multiplicities of eigenvalues) are believed to be small \((\ll_{\epsilon} (1+|\lambda|)^{\epsilon} \text{ for any } \epsilon > 0)\). However, even in that case, all that one knows in this generality about the multiplicity is an extremely modest (by a factor \(\log(2+|\lambda|)\)) improvement of Berard (see [Sar95]).

An obvious way to try to resolve the multiplicity issue is to exploit extra symmetries and to require that \(\phi\) is an eigenfunction of additional operators, commuting with \(\Delta\). For instance, if \(X\) is locally symmetric of rank \(>1\) and \(\phi\) is an eigenfunction of the full algebra of invariant differential operators the bound improves (1.1) to \(\ll (1+|\lambda|)^{(\dim X - \text{rank } X)/4}\) [Sar].

This does not help much for surfaces, but if these are of ”arithmetic type” and so endowed with the action of a suitable algebra of commuting Hecke operators, one may then consider instead joint Hecke-Laplace eigenfunctions, and some significant saving is possible. Indeed when \(X := \Gamma \backslash \mathbb{H}\) is a modular or a Shimura curve (\(\mathbb{H} \cong SL_2(\mathbb{R})/SO_2\) the hyperbolic plane and \(\Gamma \subseteq SL_2(\mathbb{R})\) an arithmetic lattice), the bound (1.1) was improved by a power of \(1+|\lambda|\) in the groundbreaking work of Iwaniec and Sarnak [IS95]. Later, a similar result was obtained by Vanderkam for \(\mathbb{H}\) replaced by the 2-sphere \(S^2 \cong SO_3(\mathbb{R})/SO_2(\mathbb{R})\) [Van97]. In these cases it follows from the multiplicity one theorem and the Jacquet-Langlands correspondence that the dimension of a Hecke-Laplace eigenspace is bounded by \(\ll_{\epsilon} \text{ vol}(X)^{\epsilon}\) for any \(\epsilon > 0\).

### 1.1. Bounds on 2-dimensional ellipsoids.

Our previous work [BM11] dealt with a family of varieties \(X = X(L,q)\) associated to pairs \((L,q)\) for \(q\) a definite quadratic form on a three dimensional \(\mathbb{Q}\)-vector space \(V\) and \(L\) a suitable lattice in \(V\); \(X\) was then a finite union of (quotients) of 2-spheres indexed by a set of representatives of genus classes of \(L\).

The present paper extends [BM11] in two further directions: on the one hand, by a new treatment of the amplifier we improve significantly the main result in [BM11]; on the other hand, we extend the argument to varieties attached to quadratic lattices \(L \subset V\) for \((V,q)\) a totally definite ternary quadratic space defined over some fixed, totally real number field \(F\) of degree \(d\) over \(\mathbb{Q}\); the corresponding variety \(X(L,V)\) is then a finite union of \(d\)-fold products of 2-spheres. We stress that this extension to number fields is not solely for the sake of generality: in the next subsection, we use these results to study similar problems for varieties associated to \textit{quaternary quadratic spaces}.

Our main results are proven under some additional assumption on the ”shape” of \(L\) which is better expressed in terms of quaternions (we refer to §2.1 for the notations related to quaternions). Given a totally real field \(F/\mathbb{Q}\) of degree \(d\), a totally definite ternary quadratic space \((V,q)\) over \(F\) and a lattice \(L \subset V\), there exists (cf. [Kit93, Chap. 1]) \(\lambda \in F^\times\) and a (unique up to isomorphism) totally definite quaternion algebra \(B\) over \(F\) (i.e. at all real places of \(F\), \(B\) is isomorphic to the Hamilton quaternions) such that \((V,q)\) is isometric to the ternary quadratic space \((B^0,\lambda nr)\); here \(B^0\) denotes the space of trace 0 quaternions and \(nr\) the reduced norm of \(B\). Choosing such an isometry, we therefore identify \(V\) with \(B^0(F)\) and \(L\) with a certain sublattice of \(B^0(F)\). We assume from now on that

\[L = \mathfrak{O}^0 := B^0(F) \cap \mathfrak{O}\]

for an Eichler order \(\mathfrak{O}\) of \(B(F)\).

To this situation is associated a finite disjoint union of quotients of \(d\)-fold products of 2-spheres

\[(1.3)\quad X^{(2)}(\mathfrak{O}) := \bigsqcup_{i \in I} X_i, \quad X_i := \Gamma_i \backslash (S^2)^d,\]

for \(\Gamma_i < SO_3(\mathbb{R})^d\) some finite subgroup of order bounded in terms of \(d\) only. This union corresponds to a double adelic quotient whose definition is given in (2.8).

The quotients \(X_i\) are called the components of \(X^{(2)}(\mathfrak{O})\) and their indexing set is the set of classes of left-\(\mathfrak{O}\) ideals in \(B\) (whose cardinality is the class number of \(\mathfrak{O}\)). Consider the restriction to \(S^2\) of the Euclidean metric on \(\mathbb{R}^3\); this is an \(SO_3(\mathbb{R})\)-invariant Riemannian metric on \(S^2\) which induces a
volume form and a $d$-tuple of Laplace operators $\Delta = (\Delta_1, \ldots, \Delta_d)$ on $(S^2)^d$ that descend to $X^{(2)}(\mathcal{O})$. One has (see §2.4 below)

$$V_2 := \text{vol}(X^{(2)}(\mathcal{O}))) = |\text{nr}_{F/Q} \text{disc}(\mathcal{O})|^{1/2+\alpha_F(1)}$$

(here the $\text{disc}(\mathcal{O})$ refers to the discriminant of the quaternion space $(\mathcal{O}, \text{nr})$).

We are interested in obtaining non-trivial bounds for the $L^\infty$-norm of an $L^2$-normalized $\Delta$-eigenfunction $\varphi$ on $X^{(2)}(\mathcal{O})$ in terms of the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_d)$ of $\Delta$ and of the total volume $\text{vol}(X^{(2)}(\mathcal{O}))$. The trivial bound in this case is (see [Sar] for a general result)

$$\|\varphi\|_\infty \ll |\lambda|^{1/4}, \text{ with } |\lambda| := \prod_{j=1}^d (1 + |\lambda_j|).$$

Our objective is to improve over this bound simultaneously in the $\lambda$ and the volume aspect; such non-trivial bounds are called “hybrid”. In this generality this is hopeless: the previous bound is indeed sharp both in the volume and in the $\lambda$-aspect. The possibility of constructing Laplace eigenvectors with large sup-norm comes from the fact that $\Delta$-eigenfunctions have very large multiplicities (roughly $\approx V_2|\lambda|^{1/2}$). As explained above, a way to resolve this issue is to require $\varphi$ to be also an eigenfunction of a family of “Hecke” operators, indexed by the complement of a finite, fixed subset of the prime ideals of $F$: $\{T_p\} \setminus \text{disc}(\mathcal{O})$. The Hecke operators $\{T_p, \frac{p}{\text{disc}(\mathcal{O})}\}$ together with $\Delta$ generate a commutative algebra of self-adjoint operators on $L^2(X^{(2)}(\mathcal{O}))$; in particular this space admits an orthonormal basis made of Laplace-Hecke eigenfunctions.

**Theorem 1.** Let $\mathcal{O}$ be an Eichler order in a totally definite quaternion algebra $B$ over $F$, and let $\varphi$ be an $L^2$-normalized Hecke-Laplace eigenfunction on $X^{(2)}(\mathcal{O})$. Then one has with the notation as in (1.5),

$$\|\varphi\|_\infty \ll |\lambda|^{\frac{1}{2}} (V_2|\lambda|^{\frac{1}{2}})^{-\frac{1}{4}}.$$  

Individually, we obtain the following bounds in the $\lambda$ and in the volume aspect

$$\|\varphi\|_\infty \ll \varepsilon |\lambda|^{\frac{1}{4}} V_2^{-\frac{1}{8}+\varepsilon}, \quad \|\varphi\|_\infty \ll |\lambda|^{\frac{1}{4}-\frac{1}{48}+\varepsilon}.$$  

The bound (1.6) is obtained by interpolation between the two bounds in (1.7). We emphasize that as in [BM11] these estimates are uniform in $B$ and $\mathcal{O}$, but we regard the number field $F$ as fixed. For the rest of the paper all implied constants may depend on $F$ as well as on a small real number $\varepsilon$ where appropriate. All other dependencies will be indicated.

The first non-trivial bound of this sort was obtained by Iwaniec and Sarnak [IS95] for $F = \mathbb{Q}$, $B$ indefinite, and a fixed order $\mathcal{O}$. For $\mathcal{O}$ varying (of square-free level) a bound simultaneously non-trivial in $\text{vol}(X^{(2)}(\mathcal{O}))$ and in $|\lambda|$ was obtained by the first named author and R. Holowinsky [BH10]. This result was extended by Templier [Tem10] to the case of a totally real number field and for $B$ indefinite at one archimedean place. In the definite case, the first non-trivial result we are aware of is due to Vanderkam [Van97]: for $F = \mathbb{Q}$, $B$ the Hamilton quaternions and $\mathcal{O}$ the maximal order, he obtained

$$\|\varphi\|_\infty \ll |\lambda|^{\frac{1}{4}} V_2^{-\frac{1}{24}+\varepsilon}.$$  

Unaware of his work, we proved in [BM11] a hybrid bound for general $B$ and any Eichler order $\mathcal{O}$ of the shape

$$\|\varphi\|_\infty \ll \varepsilon |\lambda|^{\frac{1}{4}} (V_2|\lambda|^{\frac{1}{2}})^{-\frac{1}{16}+\varepsilon},$$  

which was an interpolation between the individual bounds

$$\|\varphi\|_\infty \ll |\lambda|^{1/4} V_2^{-1/12+\varepsilon} \quad \text{and} \quad \|\varphi\|_\infty \ll |\lambda|^{11/48+\varepsilon} V_2^{1/12+\varepsilon}.$$  

\footnote{Formally speaking, the symbol $|\lambda|$ should be thought of as a single quantity, not the absolute value of some real number.}
Our present result (1.7) is stronger in both aspects. The improvement in the volume aspect comes from a new way to deal with the amplifier (occurring from the amplification method) which may be of general interest. In the $\lambda$ aspect, the improvement comes from the use of Vanderkam’s method. Our bound in (1.7), however, is marginally weaker than (1.8) because of some technical obstacles in the number field case.

We remark that the strongest conceivable result in the situation of Theorem 1 is

$$\|\varphi\|_\infty \ll |\lambda|^{1/4} (V_2 |\lambda|^{1/2})^{-1/2+\varepsilon}.$$  

It seems reasonable to conjecture that in the presently considered case of a compact manifold this bound reflects the reality, although some care has to be taken as N. Templier [Tem] has recently disproved a similar conjecture in the non-compact case.

In any case, in the level aspect, we arrive at least at 33% of the true bound which is similar to Weyl’s bound vs. the Lindelöf Hypothesis for Riemann’s zeta function. Gergely Harcos and Nicolas Templier kindly informed us that for the (indefinite) discriminant quadratic form $b^2 - 4ac$ over $\mathbb{Q}$, they obtained in [HT] the same exponent $-1/6$. This convergence of exponents obtained independently and in fairly different contexts makes it therefore likely that this result will be hard to improve with the present technology.

1.2. Application to 3-dimensional ellipsoids. We illustrate the extension of [BM11] to general totally real number fields by providing non-trivial sup-norm bounds for Hecke-Laplace eigenfunctions on manifolds $X$ that are finite unions of $d$-fold products of 3-spheres $S^3 = SO_4/\mathbb{Z}_3$ (i.e. bounds for automorphic forms of orthogonal groups in 4 variables). The main point here is that there is a close relationship between automorphic forms on orthogonal groups in 4 variables over $F$ and automorphic forms on orthogonal groups in 3-variables over a suitable (possibly split) quadratic extension $E$ of $F$ (an extension of the well known fact that $SO_4(\mathbb{R})$ is a double cover of $SO_4(\mathbb{R}) \times SO_4(\mathbb{R})$).

As above these manifolds are better described in terms of quaternion algebras: recall (see §2.2) that to any non-degenerate quadratic space $(V,Q)$ over $F$ with discriminant $\Delta$, there is canonically associated a quaternion $F$-algebra $B$, a quadratic etale $F$-algebra $E$ ($F \times F$ if $\Delta$ is a square in $F^\times$ and $F(\sqrt{\Delta})$ otherwise) and a four dimensional vector space

$$B' \subset B_E = B \otimes_F E$$

such that $(V,Q)$ is similar\(^2\) to $(B',nr)$.

Given an Eichler order $\mathcal{O} \subset B$, we associate to it the integral quadratic lattice

$$(\mathcal{O}',nr) \text{ for } \mathcal{O}' = B' \cap \mathcal{O} \otimes_{\mathcal{O}_F} \mathcal{O}_E$$

and a finite disjoint union of (quotients of) $d = [F : \mathbb{Q}]$ products of 3-spheres (cf. (6.1))

$$X^{(3)}(\mathcal{O}) = \bigsqcup_{i \in I} X_i, \quad X_i := \Gamma_i \backslash (S^3)^d,$$

where $\Gamma_i < SO_4(\mathbb{R})^d$ is finite and of order bounded in terms of $d$ only. Again the indexing set is closely related to the set of genus classes of the quaternary quadratic lattice $(\mathcal{O}',nr)$ and if one equips $S^3$ with the restriction of the Euclidean metric on $\mathbb{R}^4$, the total volume of $X^{(3)}(\mathcal{O})$ satisfies

$$V_3 := \text{vol}(X^{(3)}(\mathcal{O})) = |nr_F(\text{disc}(\mathcal{O}))|^{1+\varepsilon(1)}.$$

Similarly as above, $X^{(3)}(\mathcal{O})$ is endowed with a commutative algebra of Hecke operators commuting with the corresponding Laplace operator. For an $L^2$-normalized Hecke-Laplace eigenfunction $\varphi$ on $X^{(3)}(\mathcal{O})$ the trivial bound for its sup-norm is

$$\|\varphi\|_\infty \ll |\lambda|^{1/2},$$

(uniformly in the volume $V_3$) and we obtain here an improvement in the volume aspect:

\(^2\)Even isometric if $F$ is totally real and $q$ is positive at every archimedean place, by Eichler’s norm theorem.
Theorem 2. In the situation described above, one has
\[ \|\varphi\|_{\infty} \ll_{E,\varepsilon} |\lambda|^{1/2} V_3^{-1/6+\varepsilon} \]
for any \( \varepsilon > 0 \).

The present bound is a direct application of the arguments of the proof of Theorem 1; yet it seems to be the first instance of a non-trivial arithmetic (i.e. in the level aspect) sup-norm bound for a manifold which does not factor into surfaces. Again we obtain the same (“Weyl-type”) quality in the exponent. Several extensions are possible:

(1) We have considered here only the volume aspect. The diophantine counting Lemma 5 of §4 of this paper would yield quite directly some non-trivial hybrid bounds for some \( \text{SO}_Q \)-automorphic forms, namely those, which at each archimedean place of \( F \) correspond (via the identification \( \text{SO}_Q(F) \approx \text{SO}_4(\mathbb{R}) \)) to pure weight vectors with respect to the action of the maximal torus \( \text{SO}_2(\mathbb{R}) \times \text{SO}_2(\mathbb{R}) < \text{SO}_4(\mathbb{R}) \). Laplace-Hecke eigenfunctions on 3-dimensional ellipsoids on the other hand, correspond to \( \text{SO}_3(\mathbb{R}) \)-invariant vectors; these are potentially long linear combinations of pure weight vectors, and Lemma 5 in its present form is not sufficient to obtain hybrid bounds for such functions.

(2) The present bound depends on the quadratic extension \( E \). Making it explicit and non-trivial in this aspect requires a more precise description of the local structure of the quaternary quadratic lattices considered at the places where \( E \) is ramified and versions of the counting Lemmata 2 – 5 taking this aspect into account. Observe that in the present case, the amplification method does not a priori require that \( E \) splits at many small places (as is the case in [DFI95] or [Ven10, §7]), for the group \( \text{SO}_Q(F_v) \) has rank at least 1 for almost all places of \( F \) (the places at which \( B \) is unramified).

1.3. Organization of the paper and concluding remarks. In the next section we introduce general notations and describe how the problem translates in the adelic setting. Section 3 discusses reduction theory for totally definite quadratic forms over totally real number fields, and we discuss general results about the representation of algebraic integers by such quadratic forms in Section 4. In Section 5 we apply the pretrace formula and the amplification method in a by now standard way and reduce the problem of bounding the sup norm of Hecke-Laplace eigenforms to the diophantine problems of the previous section, that is, bounding representation numbers of quadratic forms of large discriminant of \( F \)-integral vectors that are almost parallel or almost orthogonal to a given vector. The first bound in (1.7) and the bound (1.9) follow only from Lemma 2 which is at least in principle not much more than a generalized Lipschitz principle. The second bound in (1.7) is more complicated and requires Lemmata 3 – 5. There are at least two sources of improvement compared to the analysis in [BM11]: in the present paper we use explicitly the fact that the considered quadratic forms are associated to an order in a quaternion algebra and in particular represent 1. Moreover, we exploit the average over the amplifier and treat the quadratic part of variable \( \ell \) in the amplifier essentially as a new variable of the quadratic form.

Finally we would like to thank the referee for his unusually careful reading of the manuscript and a long list of constructive suggestions that greatly improved the presentation.

2. Preliminaries

Let \( F/\mathbb{Q} \) be a totally real number field of degree \( d \), \( \mathcal{O}_F \) its ring of integers, \( U = \mathcal{O}_F^\times \) its group of units and \( U^+ \) the subgroup of totally positive units. For a place \( v \) of \( F \), we denote by \( F_v \) the associated local field. A typical real place of \( F \) will be denoted as an embedding
\[ \sigma : F \hookrightarrow \mathbb{R} \]
and the list of real places will be denoted by \( \sigma_1, \ldots, \sigma_d \); for \( x \in F \), we write \( x^\sigma = \sigma(x) \in \mathbb{R} \) for the corresponding conjugate.
We denote by \( A = \coprod_v F_v \), \( A_f = \coprod_{v < \infty} F_v \), \( F_\infty = \prod_v F_v \) the \( F \)-algebras of adèles, finite adèles and archimedean components of \( F \). We denote the norm on \( F/\mathbb{Q} \) by \( \text{nr}_F \) and use the same notation for the natural extension of the norm to the \( F \)-ideals or to various \( F \)-algebras related to \( F \) (\( F_v, F_\infty, A \) etc.). To ease notations, for \( \mathcal{I} \subset F \) an \( F \)-ideal or \( x_f \in \mathcal{I}_f \) a finite \( F \)-idèle we will freely identify \( \text{nr}_F(\mathcal{I}) \) or \( \text{nr}_F(x_f) \) with the positive generator of its underlying \( \mathbb{Z} \)-ideal in \( \mathbb{Q} \), so that if necessary the expression \( \text{nr}_F(\mathcal{I})^{\sqrt{\pi}} \) (say) is well-defined.

### 2.1. Ternary quadratic spaces and quaternion algebras.

We recall some facts about quaternion algebras, see e.g. [Vig80] or [BM11] for more details. Let \( B \) be a totally definite quaternion algebra defined over \( F \). We denote its canonical involution by \( z \mapsto z^* \) and its reduced trace and reduced norm by

\[
\text{tr} : z \mapsto z + z^*, \quad \text{nr} : z \mapsto zz^*.
\]

We denote the trace-0 quaternions and trace-0 quaternions of norm 1 by \( B^0 \) and \( B^{0,1} \) respectively (considered as algebraic varieties over \( F \)). The spaces \( (B^0, \text{nr}), (B, \text{nr}) \) are quadratic \( F \)-spaces whose associated inner product is denoted by

\[
\langle z_1, z_2 \rangle_B = \frac{1}{2} \text{tr}(z_1 z_2^*), \quad \langle z, z \rangle_B = \text{nr}(z).
\]

We denote by \( B^\times \) the group of units, \( Z \) its center (the subgroup of scalars), \( B^1 \) the subgroup of quaternions of reduced norm 1 and by \( \text{PB}^\times = Z \backslash B^\times \) the projective quaternions. All these are considered as \( F \)-algebraic groups in the evident way. We write \( B(A), B^\times(A), B^1(A), B^\times(A_f), \ldots, B(F_v), B^\times(F_v) \) etc. for the sets of rational points of these varieties over the corresponding \( F \)-algebras.

The conjugation action of the group of units \( B^\times \) on the ternary quadratic space \( (B^0, \text{nr}) \) is isometric (i.e. preserves the norm form) and the map

\[
g \in B^\times \mapsto \rho_g : B^0 \times \mathbb{F}_v \to B^0 \times \mathbb{F}_v \quad x \mapsto gxg^{-1}
\]

is an isomorphism of \( F \)-algebraic groups

\[
\text{PB}^\times = Z \backslash B^\times \cong \text{SO}(B^0), \quad B^1 \cong \text{Spin}(B^0)
\]

where \( \text{Spin}(B^0) \) denotes the spin group (the simply connected covering group of \( \text{SO}(B^0) \)).

A place \( v \) is called ramified if \( B_v := B \otimes_F F_v \) is a division algebra, and non-ramified otherwise; in the former case, \( B_v \) is the unique (up to isomorphism) quaternion division algebra over \( F_v \); in the latter \( B_v \cong \text{Mat}(2,F_v) \) in which case the reduced norm and reduced trace are given by the usual determinant and trace for matrices. Since \( B \) is totally positive, all archimedean places are ramified. The reduced discriminant \( D_B \) of \( B \) is the product of the finite ramified prime ideals.

A lattice or ideal \( \mathcal{I} \subset B \) is an \( \mathcal{O}_F \)-module of maximal rank 4. An order \( \mathcal{O} \) is a subring of \( B \) which is also a lattice. The left order \( \mathcal{O}_l(\mathcal{I}) \) of a lattice \( \mathcal{I} \) is the set \( \mathcal{O}_l(\mathcal{I}) = \{ \gamma \in B, \gamma \mathcal{I} \subset \mathcal{I} \} \). Given an order \( \mathcal{O} \), a left \( \mathcal{O} \)-ideal is defined as a lattice \( \mathcal{I} \) such that \( \mathcal{O}_l(\mathcal{I}) = \mathcal{O} \). Two left \( \mathcal{O} \)-ideals \( \mathcal{I}, \mathcal{I}' \) are called (right-)equivalent if there exists \( \gamma \in B^\times(F) \) such that \( \mathcal{I}' = \gamma \mathcal{I} \). The set of such equivalence classes is denoted \( \text{Cl}(\mathcal{O}) \); this set is finite and its cardinality \( h(\mathcal{O}) = |\text{Cl}(\mathcal{O})| \) is the (left ideal) class number of \( \mathcal{O} \). The discriminant of an order \( \mathcal{O} \) is by definition \( \text{disc}(\mathcal{O}) = \text{det}(\text{tr}(\gamma \Theta_{i,j})_{i,j}), \) for an \( \mathcal{O}_F \)-basis \( \{\gamma_1, \ldots, \gamma_4\} \) of \( \mathcal{O} \). The reduced norm \( \text{nr}(\mathcal{I}) \) of a lattice \( \mathcal{I} \) is the fractional \( \mathcal{O}_F \)-ideal generated by all elements \( \text{nr}(\gamma) \) with \( \gamma \in \mathcal{I} \). The dual of a lattice \( \mathcal{I} \) is the lattice \( \mathcal{I}^* = \{ \gamma \in B(F), \text{tr}(\gamma \mathcal{I}) \subset \mathcal{O}_F \} \).

One defines the reduced discriminant \( \text{disc}^*(\mathcal{O}) \) of \( \mathcal{O} \) to be the ideal

\[
\text{disc}^*(\mathcal{O}) := \text{nr}((\mathcal{O}^*)^{-1}) = \text{nr}(\mathcal{O}^*)^{-1},
\]

so the reduced discriminant is the level of the associated norm form. If \( \mathcal{O} \) is a maximal order, its reduced discriminant equals the reduced discriminant \( D_B \) of \( B \) [Vig80, II.4.7]. In general, one has the following important relation between discriminant and reduced discriminant [Vig80, I.4.7]

\[
\text{disc}(\mathcal{O}) = \text{disc}^*(\mathcal{O})^2.
\]
An Eichler order is by definition the intersection of two maximal orders. To an Eichler order there is associated an $\mathcal{O}_F$-ideal $\mathfrak{N}$ coprime to $D_B$ such that for every $p$ coprime with $D_B$, $\mathcal{O}_p$ is conjugate to the order

\[
(2.3) \quad \left( \mathcal{O}_{F,p}, \mathfrak{N} \mathcal{O}_{F,p}, \mathcal{O}_{F,p} \right) \subset M_2(F_p)
\]

where $\mathfrak{N}$ is a finite idele corresponding to the ideal $\mathfrak{N}$. We may and will assume that the choice of $\mathcal{O}_{\text{max}}$ is such that $\mathcal{O}_p$ corresponds precisely to (2.3). Eichler orders associated with the same $\mathfrak{N}$ are locally conjugate (and conversely). Moreover, the left order $\mathcal{O}(I)$ of a right $\mathcal{O}$-ideal $I$ is everywhere locally conjugate to $\mathcal{O}$, hence the norm forms of $\mathcal{O}(I)$ and $\mathcal{O}$ are in the same genus.

For an Eichler order $\mathcal{O}$, the discriminant and the reduced discriminant have the following explicit expressions [Vig80, p. 85]

\[
disc^*(\mathcal{O}) = D_B \mathfrak{N}, \quad \text{disc}(\mathcal{O}) = (D_B \mathfrak{N})^2,
\]

and the class number equals [Vig80, p. 143]

\[
h(\mathcal{O}) \asymp \text{nr}_F(D_B \mathfrak{N}) \prod_{p|\mathfrak{D}_B} \left(1 - \frac{1}{\text{nr}_F(p)}\right) \prod_{p|\mathfrak{N}} \left(1 - \frac{1}{\text{nr}_E(p)}\right)^{-1} = \text{nr}_F(\text{disc}(\mathcal{O}))^{1/2 + o(1)}.
\]

2.2. Quaternary quadratic spaces and quaternion algebras. Let $(V, Q)$ be a non-degenerate quaternary quadratic space over $F$ and let $E$ be the quadratic $F$-algebra

\[
E = \begin{cases} F \times F & \text{if disc}(Q) \in (F^\times)^2 \text{ (i.e. is a square class),} \\ F(\sqrt{\text{disc}(Q)}) & \text{if disc}(Q) \notin (F^\times)^2,} \end{cases}
\]

equipped with either the $F$-invariant involution $\sigma(x, y) = (y, x)$ if $E = F \times F$ or the canonical $F$-invariant involution if $E$ is a field. In the split case we view $F$ as embedded diagonally into $F \times F$.

There is a unique quaternion algebra $B$ defined over $F$ such that $(V, Q)$ is similar to the quaternion quadratic space $(B', \text{nr})$ that we now describe (see also [Pon76]). Let $B_E := B \otimes_F E$ (this is often referred to as the second Clifford algebra of the quadratic space $V$). Slightly abusing notations, we denote by $\cdot^*$ the extension to $B_E$ of the canonical involution of $B$, by $\text{nr}(z) = zz^*$ the associated norm form on $B_E$ and by $\sigma = \text{Id}_B \otimes \sigma$ the extension of $\sigma$ from $E$ to $B_E$. Let

\[
B' := \{z \in B_E, \ \sigma(z)^* = z\}.
\]

Then $\text{nr}$ is $F$-valued on $B'$ and $(B', \text{nr})$ defines a non-degenerate quaternary quadratic space over $F$ such that $\text{disc}(B')$ is a square if $E = F \times F$ and $E = F(\sqrt{\text{disc}(B')})$ otherwise.

We now proceed to describe the orthogonal group $\text{SO}(B')$ along these lines: for any $w \in B_E^\times$, the map

\[
z \mapsto wz\sigma(z)^*
\]

leaves $B'$ invariant and defines a proper similitude with factor $\lambda(w) = \text{nr}_{E/F}(\text{nr}(w))$. In particular, if $w$ is such that $\text{nr}(w) \in F$ (i.e. $\text{nr}(w) = \sigma(\text{nr}(w))$), the map

\[
\rho_w : z \in B' \mapsto wz\sigma(z)^{-1} = \frac{1}{\text{nr}(w)}wz\sigma(w)^*
\]

is a special orthogonal transformation of $(B', \text{nr})$; moreover the map $w \mapsto \rho_w$ induces an isomorphism of $F$-algebraic groups

\[
\text{SO}(B') \cong Z_F \setminus \{w \in B_E^\times, \ \text{nr}(w) = \sigma(\text{nr}(w))\}.
\]

Here we view $B_E^\times$ as an $F$-algebraic group (of dimension 8) and $\text{nr} : B_E^\times \to \mathbb{G}_m, E$ and $\sigma : B_E^\times \to B_E^\times$ as algebraic maps. We also note that the stabilizer in $\text{SO}(B')$ of the vector $1 \in B'$ is precisely

\[
\text{SO}(B')_1 = Z_F \setminus \{w \in B_E^\times\} = PB_E^\times = \text{SO}(B_F^0).
\]
where $B'^θ = B_E^θ \cap B'$ is the orthogonal subspace to 1. In the split case we have $B_E^θ = B^θ \times B^θ$, $B'$ is identified with $B$ via the embedding $z \in B \mapsto (z, z^*) \in B \times B$, and the identification

$$\text{SO}(B) \simeq Z_F \setminus \{(w, w') \in B_E^θ \times B_E^θ, \ \nr(w) = \nr(w')\}$$

($Z$ diagonally embedded in $B^θ \times B^θ$) is given via the map

$$(w, w') \mapsto (\rho_{w,w} : z \mapsto wzw'^{-1}).$$

We denote by $G < \tilde{G}$ the $F$-algebraic groups defined (at the level of their $F$-points) by

$$(2.4) \quad G = Z_F \setminus \{w \in B_E^θ, \ \nr(w) = \sigma(\nr(w))\} \simeq \text{SO}(B'), \quad \tilde{G} = Z_F \setminus B_E^θ.$$

2.3. Representations at the archimedean place. For any integer $m \geq 0$ there is a unique irreducible (unitary) representation of $\text{SU}_2(\mathbb{C}) \simeq \text{Spin}_3(\mathbb{R})$ of degree $d_m = m + 1$, denoted $\pi_m$, and any irreducible representation of $\text{SU}_2(\mathbb{C})$ is isomorphic to some $\pi_m$, see e.g. [Far08, Section 7] for details. The representation $\pi_m$ may be realized concretely as the space of complex homogeneous polynomials of degree $m$ in two variables on which $\text{SU}_2(\mathbb{C}) \subset \text{GL}_2(\mathbb{C})$ acts by linear change of variables. The Casimir element (say with respect to the inner product on the Lie algebra $(X,Y) = -\frac{1}{2}\text{Tr}(XY)$) $C_{\text{SU}_2(\mathbb{C})} \in U(\mathfrak{su}_2)\mathbb{C}$ acts on any realization by multiplication by the scalar

$$\lambda_m = -m(m + 2).$$

Let $T(\mathbb{R}) \cong \text{SO}_2(\mathbb{R}) \subset \text{SU}_2(\mathbb{C})$ be the stabilizer of (say) the north pole of $S^2$ under the natural projection $\text{SU}_2(\mathbb{C}) \to \text{SO}_3(\mathbb{R})$. This is a maximal torus of $\text{SU}_2(\mathbb{C})$ isomorphic to $\mathbb{R}/2\pi\mathbb{Z}$ and whose image in $\text{SO}_3(\mathbb{R})$ is the group of matrices

$$\text{SO}_2(\mathbb{R}) = \left\{ \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) & 0 \\ \sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\}.$$

Let $\theta \in \mathbb{R}/2\pi\mathbb{Z} \to \kappa(\theta) \in T(\mathbb{R})$ be a parametrization and $e : T(\mathbb{R}) \to \mathbb{C}^1$ be the character

$$e(\kappa(\theta)) = \exp(i\theta).$$

If $V_m$ is any vector space realizing $\pi_m$ and $l \in \mathbb{Z}$, let $V^l_m$ be the subspace of vectors “of weight $l$”, that is, the vectors satisfying

$$\varphi \in V_m, \quad \kappa.\varphi = e(\kappa)^l\varphi \quad \text{for all } \kappa \in T(\mathbb{R}).$$

Then $V^l_m$ is one dimensional if $|l| \leq m$, $l \equiv m \ (2)$ and zero otherwise.

**Remark 2.1.** The representation $\pi_m$ occurs in the right regular representation $L^2(\text{SU}_2(\mathbb{C}))$ with multiplicity $d_m$. When $m$ is even, $\pi_m$ descends via the natural projection $\text{SU}_2(\mathbb{C}) \to \text{SO}_3(\mathbb{R})$ to an irreducible representation of $\text{SO}_3(\mathbb{R})$. The direct sum of the weight zero vectors of each such copy of $\pi_m$ therefore injects into $L^2(\text{SO}_3(\mathbb{R}))^{\text{SO}_3(\mathbb{R})} = L^2(S^2)$ and the image is the space of harmonic homogenious polynomials of degree $m/2$ in $\mathbb{R}^3$ (i.e. polynomials $P$ such that $\Delta_\mathbb{R}^3 P = 0$). The action of the Casimir element on this space corresponds to that of a fixed multiple of the Laplace operator $\Delta_{S^2}$.

Given a non-zero $\varphi \in V^l_m$ and $g \in \text{SU}_2(\mathbb{C})$, we write

$$p_{m,l}(g) := \frac{(\varphi, g.\varphi)_m}{(\varphi, \varphi)_m}$$

for the corresponding normalized matrix coefficient, where $( \ , \ )_m$ is some $\text{SU}_2(\mathbb{C})$-invariant inner product on $V_m$. By definition $g \mapsto |p_{m,l}(g)|$ is bi-$T(\mathbb{R})$-invariant, and therefore depends only on $t = t(g) = (\rho_g(x_3), x_3)|_{\mathbb{R}^3} \in [-1, 1]$, the inner product of the north pole $x_3$ on $S^2$ with its image by the corresponding rotation. The following decay estimate holds as $g \in \text{SU}_2(\mathbb{C})$ gets “away” from $T(\mathbb{R})$ (i.e. $t$ gets away from $\pm 1$). We will use it for the bound in the eigenvalue aspect in Section 5.3.
Lemma 1. For $-m \leq l \leq m$ one has

$$|p_{m,l}(g)| \ll \min \left(1, \left(\frac{m+1}{|l|+1}\right)^{-1/2} (1-t^2)^{-1/4}\right).$$

Proof. By symmetry we may assume $0 \leq l \leq m$. Being a matrix coefficient, it is clear that $|p_{m,l}(g)| \leq 1$. One has [VK91, §6.3.1 & 6.3.7]

$$|p_{m,l}(g)| = |P^{m}_{l/2}(t)| = \left(\frac{1+t}{2}\right)^l |P^{0,2l}_{m-l}(t)|$$

where (for $\alpha, \beta \geq 0$ integers)

$$P^{(\alpha,\beta)}(t) = \frac{(-1)^n}{2^n n!} (1-t)^{-\alpha} (1+t)^{-\beta} \frac{d^n}{dt^n} [(1-t)^\alpha (1+t)^\beta (1-t^2)^n]$$

is the Jacobi polynomial. Let us recall that $P^{(\alpha,\beta)}_n$ has degree $n$ and that $\{P^{(\alpha,\beta)}_n \mid n \geq 0\}$ is orthogonal with respect to the inner product

$$\langle P,Q \rangle_{(\alpha,\beta)} := \int_{-1}^1 P(t)\overline{Q(t)} (1-t)^{\alpha} (1+t)^{\beta} dt,$$

and that [VK91, 6.10.1(7)]

$$\langle P^{(\alpha,\beta)}_n, P^{(\alpha,\beta)}_m \rangle_{(\alpha,\beta)} = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1} \frac{(n+\alpha)!(n+\beta)!}{(n+\alpha+\beta)!n!}.$$

In particular,

$$\langle P^{(0,2l)}_{m-l}, P^{(0,2l)}_{m-l} \rangle_{(0,2l)} = \frac{2^{2l+1}}{2m+1} \frac{(m-l)!(m+l)!}{(m+l)!} = \frac{2^{2l+1}}{2m+1},$$

By [EMN94, Thm. 1] we have

$$(1-t)^{\alpha/2} (1+t)^{\beta/2} P^{(\alpha,\beta)}_n(t) \ll \langle P^{(\alpha,\beta)}_n, P^{(\alpha,\beta)}_n \rangle_{(\alpha,\beta)}^{1/2} \frac{(\alpha+\beta+1)^{1/2}}{(1-t^2)^{1/4}}$$

for $t \in [-1,1]$ and hence

$$(1+t)^{l+1} P^{(0,2l)}_{m-l}(t) \ll \frac{1}{(1-t^2)^{1/4} (2m+1)^{1/2}}.$$

Remark 2.2. The above bound exhibits significant decay as $t$ gets away from $\pm 1$ uniformly for $|l| \leq (1+m)^{1-\delta}$ for any fixed $\delta > 0$. It is plausible that this holds also for very large values of $l \gg m$: for instance in the extreme case $l = m$ one has:

$$|p_{m,m}(g)| = \left(\frac{1+m}{2}\right)^m.$$

More generally it is conjectured in [EMN94] that in (2.5), the term $\frac{m+1}{|l|+1}$ can be replaced by $\frac{m+1}{\sqrt{|l|+1}}$.

We extend these notation to irreducible representations of products: for $m = (m_{\sigma})_\sigma$ a $d$-tuple of non-negative integers, we set

$$\pi_m = \bigotimes_{\sigma=1}^d \pi_{m_{\sigma}} \in \text{Irr}(SU_2(\mathbb{C})^d)$$

and denote by $d_m = \prod_{\sigma} (m_{\sigma}+1) = |m|$ its dimension (cf. footnote 1). Given a realization $V_m = \bigotimes_{\sigma} V_{m_{\sigma}}$ of $\pi_m$ and a $d$-tuple $l = (l_{\sigma})_\sigma$ of integers, we denote by $V^l_m = \bigotimes_{\sigma} V^l_{m_{\sigma}}$ the tensor product of weight $l_{\sigma}$ vectors with respect to the product $T(\mathbb{R})^d$, and for $g_\infty = (g_{\sigma})_{\sigma} \in SU_2(\mathbb{C})^d$ we denote by

$$p_{m,l}(g) := \prod_{\sigma} p_{m_{\sigma},l_{\sigma}}(g_{\sigma})$$

the corresponding normalized matrix coefficient.
2.4. Adelic interpretation of ellipsoids. As in [BM11, §4], we now define $X^{(2)}(\mathcal{O})$ as an adelic quotient. We refer to [Vig80, Chap. III & V] for more details.

2.4.1. Archimedean place. Fix once and for all $\{x_1, x_2, x_3\} = \{x_1, x_2, x_3, x_3, x_3\}_\sigma$ an orthonormal basis of $B^0(F_\infty) = \prod_\sigma B^0(F_\sigma) \simeq (\mathbb{R}^3)^d$ for the quadratic form $\oplus_\sigma \text{nr};$ this induces identifications

$$B^1(F_\infty) \simeq \text{SU}_2(\mathbb{C})^d, \quad PB^\times(F_\infty) \simeq \text{SO}_3(\mathbb{R})^d.$$ 

Let

$$x^0 := x_3 = (x_3, \sigma) \in B^{0,1}(F_\infty) \simeq (S^2)^d \text{ and } K_\infty \simeq \text{SO}_2(\mathbb{R})^d$$

be its stabilizer under the conjugacy action of $B^\times$. This yields an identification

$$(2.7) \quad Z(F_\infty) \backslash B^\times(F_\infty)/K_\infty \simeq (S^2)^d.$$ 

2.4.2. Finite places. Given an Eichler order $\mathcal{O}$ (the intersection of two maximal orders), we denote by $\hat{\mathcal{O}}$ the closure of $\mathcal{O}$ in $B(\hat{k}_f).$ Correspondingly let $K_f = \hat{\mathcal{O}}^\times \subset B^\times(\hat{k}_f)$ be the subgroup of units of the ring $\hat{\mathcal{O}}$ ($K_f$ is an open compact subgroup of $B^\times(\hat{k}_f)$) and $\hat{\mathcal{O}}^0 \subset B^0(\hat{k}_f)$ the elements of $\hat{\mathcal{O}}$ with trace 0.

We define the manifold in Theorem 1 as the adelic quotient

$$(2.8) \quad X^{(2)}(\mathcal{O}) := Z(\hat{k}_f)B^\times(F)\backslash B^\times(\hat{k}_f)/K_\infty \hat{\mathcal{O}}^\times.$$ 

Because of (2.7), $X^{(2)}(\mathcal{O})$ is therefore identified (cf. (1.3)) with a finite disjoint union of quotients of $d$-fold product of spheres indexed by the finite double coset $Z(\hat{k}_f)B^\times(F)\backslash B^\times(\hat{k}_f)/\hat{\mathcal{O}}^\times$. Alternatively one could also think of $X^{(2)}(\mathcal{O})$ as a collection of quotients of ellipsoids associated to different quadratic forms (covering all classes in the genus of $(\mathcal{O}^0, \text{nr})$).

Moreover, if $\mathcal{O}$ is assumed to be an Eichler order, every $\mathcal{O}$-ideal in $B(F)$ is locally principal and the double cosets $Z(\hat{k}_f)B^\times(F)\backslash B^\times(\hat{k}_f)/\hat{\mathcal{O}}^\times$ parametrize precisely $\text{Cl}(\mathcal{O})$.

2.4.3. Measures. We set

$$[B^\times] := Z(\hat{k}_f)B^\times(F)\backslash B^\times(\hat{k}_f);$$

this quotient (equipped with the quotient of Haar measures, see [Vig80, Cor. 2.3, Chap. V]) has finite volume which is equal to 2 when the chosen measures are the Tamagawa measures. The volume of $\hat{\mathcal{O}}^\times \subset B^\times(\hat{k}_f)$ for the Tamagawa measure is of size $\asymp F^{n_F} \text{nr}_F(\text{disc}(\mathcal{O}))^{-1/2+\sigma(1)}$; therefore, if we normalize Haar measures on $B^\times$ so as to match the normalization given in the introduction at the infinite places, and for the finite places so that $\hat{\mathcal{O}}^\times$ has mass 1, and denote the resulting measure on $[B^\times]$ by $dg$, we obtain

$$\int_{[B^\times]} dg = \text{vol}(X^{(2)}(\mathcal{O})) = V_2$$

which is precisely (1.4). The corresponding inner product will be denoted by

$$\langle \varphi, \varphi \rangle = \int_{[B^\times]} |\varphi(g)|^2 dg.$$ 

2.4.4. Automorphic forms. Our original problem is equivalent to bounding a certain $L^2$-normalized automorphic function $\varphi$ on the adelic quotient $Z(\hat{k}_f)B^\times(F)\backslash B^\times(\hat{k}_f)$, which is

1. right $K_\infty,K_f$-invariant,
2. an eigenfunction of the Casimir operators $C = (C_\sigma)_\sigma$ associated to the group $B^1(F_\infty) \simeq \text{Spin}_4(\mathbb{R})^d$ with eigenvalues

$$\lambda_\sigma = -m_\sigma(m_\sigma + 2)$$

for a $d$-tuple $m$ of even integers as in the preceding subsection,
3. an eigenfunction of a certain Hecke algebra $H(\mathcal{O})$, which is a commutative algebra of normal operators commuting with the Casimir operators $C.$
We will recall the definition of the Hecke algebra \( \mathcal{H}(\mathcal{O}) \) in section §5; in particular we do not consider Hecke operators at ramified primes, so that the operators of \( \mathcal{H}(\mathcal{O}) \) are indeed normal. Our assumptions imply that the \( B^\times(\mathcal{A}) \)-translates of \( \varphi \) generate an automorphic representation \( \pi = \otimes_v \pi_v \) of \( B^\times(\mathcal{A}) \) with trivial central character such that

\[
\pi_\infty = \otimes_v \pi_v \simeq \pi_m
\]

for \( m = (m_\sigma) \) given in (2.9) and \( \pi_m \) defined in §2.3 (under the identifications of §2.4.1), and such that \( \pi_v \) is an unramified representation for every finite place \( v \notmid \text{disc}(\mathcal{O}) \).

If \( \pi \) is finite-dimensional, then it is one-dimensional and \( \varphi \) is proportional to the function

\[
g \in B^\times(\mathcal{A}) \to \chi(\text{nr}(g))
\]

for some (quadratic) character \( \chi \) on \( F^\times \backslash \mathbb{A}^\times \). In this case \( \varphi \) is constant on the various components of \( X^{(2)}(\mathcal{O}) \) with value equal to \( \pm V_2^{-1/2} \). In particular its Laplace eigenvalues are \( (0, \cdots, 0) \), and the bounds of Theorem 1 are a fortiori satisfied. Therefore we can restrict ourselves to infinite dimensional representations for the rest of the paper.

2.5. A slight generalization. With no extra effort we can consider a slightly more general setting: let \( \chi : F^\times \backslash \mathbb{A}^\times \to \mathbb{C}^\times \) be a unitary Hecke character, and \( L^2(Z(\mathcal{A})B^\times(F)\backslash B^\times(\mathcal{A}), \chi) \) the space functions on \( B^\times(F)\backslash B^\times(\mathcal{A}) \) satisfying

\[
\varphi(\lambda \gamma g) = \chi(\lambda)\varphi(g), \quad \lambda \in Z(\mathcal{A}), \quad \gamma \in B^\times(F), \quad g \in B^\times(\mathcal{A})
\]

and such that \( \langle \varphi, \varphi \rangle = \int_{B^\times(\mathcal{A})} |\varphi(g)|^2 \, dg \) is finite.

Let \( \pi \simeq \otimes_v \pi_v \to L^2(Z(\mathcal{A})B^\times(F)\backslash B^\times(\mathcal{A}), \chi) \) be an infinite dimensional irreducible automorphic representation of \( B^\times(\mathcal{A}) \) with central character \( \chi \). Under the identification \( B^1(F_\infty) \simeq SU_2(\mathbb{C})^d \), the representation \( \pi_\infty := \otimes_{\sigma | \infty} \pi_{\sigma} \) corresponds to \( \pi_m \) for some \( m \in \mathbb{N}_0^d \). Let

\[
\lambda = (\lambda_\sigma), \quad \lambda_\sigma = -m_\sigma(m_\sigma + 2)
\]

be the eigenvalues of the Casimir operators \( (C_{B^1(F_\sigma)})_\sigma \) and let as in (1.5)

\[
|\lambda| = \prod_\sigma (1 + |\lambda_\sigma|).
\]

Let \( \mathcal{O} \subset B \) be an Eichler order and let \( \varphi \in \pi \) be a non-zero smooth, \( \hat{\mathcal{O}}^\times \)-invariant function of some fixed weight \( l = (l_\sigma) \in \mathbb{Z}^d \) with respect to action of the maximal torus \( K_\infty \simeq SO_2(\mathbb{R})^d \) (in particular \( \pi_v \) is an unramified principal series representation at every place not dividing \( \text{disc}(\mathcal{O}) \) and \( \chi \) is unramified at these places as well).

Under these conditions, we prove the following slightly more general version of Theorem 1:

**Theorem 3.** Let \( \mathcal{O} \) be an Eichler order in a totally definite quaternion algebra \( B \) over \( F \), and let \( \varphi \in \pi \subseteq L^2(Z(\mathcal{A})B^\times(F)\backslash B^\times(\mathcal{A}), \chi) \) as described in this subsection. Then one has

\[
\|\varphi\|_\infty \ll_{1} |\lambda|^{1/2} (V_2|\lambda|^{1/2})^{-\frac{d}{2}} \|\varphi\|_2.
\]

We obtain the following individual bounds in the \( \lambda \) and in the volume aspect

\[
\|\varphi\|_\infty \ll_{1} |\lambda|^{1/2} V_2^{-1/2} \varepsilon, \quad \|\varphi\|_\infty \ll_{1} |\lambda|^{1/2 - \frac{d}{2}} \|\varphi\|_2.
\]

All implied constants depend at most on \( F, \varepsilon \) and \( l \).

The next three sections are devoted to the proof of this theorem.

**Remark 2.9.** Here we have assumed that the weight \( l = (l_\sigma) \) of \( \varphi \) is fixed. This is merely to simplify exposition (and also because the lowest weight case is arguably the most interesting one). The proof of Theorem 3 together with the bound (2.5) yields immediately a non-trivial bound for \( \|\varphi\|_\infty \) uniformly across all weights \( l \) satisfying

\[
|l| = \prod_\sigma (1 + |l_\sigma|) \leq |m|^{1-\delta}
\]
for any fixed $\delta > 0$. We expect that such a non-trivial bound holds for all $1$, and this would follow from good enough bounds for Jacobi polynomials (cf. Remark 2.2).

3. Reduction of definite quadratic forms

Let

$$Q(x) = \frac{1}{2} \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j, \quad a_{ij} = a_{ji} \in \mathcal{O}_F, \quad a_{jj} \in 2\mathcal{O}_F,$$

be an $F$-integral quadratic form in $n$ variables. Let $A = (a_{ij})_{1 \leq i, j \leq n} \in \text{Mat}_{n \times n}(\mathcal{O}_F)$ be the symmetric $n \times n$-matrix associated to $Q$. The determinant of $Q$ is

$$\Delta = \det A.$$

Note that the determinants of two equivalent forms over $\mathcal{O}_F$ may differ by the square of a unit. The quadratic form defines a bilinear form

$$(3.1) \quad \langle x, y \rangle := \frac{1}{2} x^t A y = \frac{1}{2} (Q(x + y) - Q(x) - Q(y)).$$

Let $\mathcal{O}_F':= \{y \in F^n \mid 2(y, y_F) \subseteq \mathcal{O}_F\}$. The level of $Q$ is the integral ideal $n := (Q(\mathcal{O}_F')\mathcal{O}_F)^{-1}$. In particular, if $N \in n$, then $NA^{-1} \in \text{Mat}_{n \times n}(\mathcal{O}_F)$ is an integral matrix. (Indeed, if $N \in n$, then by definition $Nx^t A x \in 2\mathcal{O}_F$ for all $x \in F^n$ such that $x^t A y \in \mathcal{O}_F$ for all $y \in \mathcal{O}_F'$, hence $z^t (NA^{-1}) z \in 2\mathcal{O}_F$ for all $z \in \mathcal{O}_F'$, which implies that $NA^{-1}$ is integral.)

For any real embedding $\sigma : F \hookrightarrow \mathbb{R}$ denote by $Q^\sigma$ the conjugated form. We assume that $Q$ is totally positive definite, that is, $Q^\sigma$ is positive definite for all $\sigma$.

Minkowski developed a reduction theory for rational positive definite quadratic forms (see e.g. [Cas78, Chapter 12]) that has been extended to arbitrary number fields by Humbert [Hum40]. We summarize some basic facts. Every quadratic form is equivalent (over $\mathcal{O}_F$) to some form of the shape

$$(3.2) \quad Q(x) = \frac{1}{2} \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j = h_1(x_1 + c_{12} x_2 + \ldots + c_{1n} x_n)^2 + h_2(x_2 + c_{23} x_3 + \ldots + c_{2n} x_n)^2 + \ldots + h_n x_n^2,$$

with $c_j, h_j \in F$ where

$$(3.3) \quad a_{ij}^\sigma \ll a_{jj}^\sigma \asymp h_j^\sigma$$

for all $1 \leq i, j \leq n$ and all embeddings $\sigma$, and

$$(3.4) \quad 1 \ll h_1^\sigma \times \ldots \times h_n^\sigma \ll h_2^\sigma \times \ldots \times h_n^\sigma \ll \ldots \ll h_1^\sigma \times \ldots \times h_n^\sigma.$$

Here and henceforth all implied constants depend only on $n$ and $F$. Clearly,

$$(3.5) \quad h_1^\sigma \cdot \ldots \cdot h_n^\sigma = \Delta^\sigma$$

where (by slight abuse of notation) $\Delta$ is the determinant of the form (3.2). This determinant (which may differ from the determinant of the original form $Q$) has the advantage that its conjugates are of comparable size. From now on we will always refer to this balanced determinant when we use the symbol $\Delta$. Of course, this convention is not necessary when we use the norm of $\Delta$.

We denote the eigenvalues of the matrix $A^\sigma = (a_{ij}^\sigma)$ by $0 < \lambda_1^\sigma \leq \lambda_2^\sigma \ldots \leq \lambda_n^\sigma$. By (3.3) – (3.5), the determinant of any $(n - 1) \times (n - 1)$ submatrix of $A^\sigma$ is $O(\Delta^\sigma)$, hence by Cramer’s rule the eigenvalues of $(A^\sigma)^{-1}$ are $O(1)$, and therefore

$$(3.6) \quad 1 \ll \lambda_1^\sigma \leq \lambda_n^\sigma \ll \Delta^\sigma.$$

Let $\tilde{Q}$ be the quadratic form in $n-1$ variables that is derived from $Q$ by setting $x_n = 0$. Let $\tilde{A}$ be the corresponding $(n-1) \times (n-1)$-submatrix of $A$, and denote by $\tilde{\Delta}$ its determinant. The $(n, n)$th-entry
of \( A^{-1} \) is by Cramer’s rule \( \hat{\Delta}/\Delta \) (up to sign); hence \( N\hat{\Delta}/\Delta \) is integral for all \( N \in \mathfrak{n} \). Therefore the ideal \( \mathfrak{n}(\hat{\Delta})/(\Delta) \) is integral, and we obtain

\[
\text{nr}_F(h_1) \times \ldots \times \text{nr}_F(h_{n-1}) = \text{nr}_F(\hat{\Delta}) \geq \frac{\text{nr}_F(\Delta)}{\text{nr}_F(\mathfrak{n})}
\]

(3.7)

4. Representation numbers of quadratic forms

In this section we establish several lemmata to bound certain averages of representation numbers of \( \mathcal{O}_F \)-integers by some totally definite quadratic form \( Q \). To perform the counting we will frequently use the following consequence of Dirichlet’s unit theorem: let \( A_1, \ldots, A_d > 0 \) be any positive real numbers and write \( A = A_1 \ldots A_d \). Then

\[
\# \{ u \in U : |u^\sigma| \leq A_j \} \ll_F (2 + A)^{d-1}.
\]

As a consequence we find

(4.1)

\[
\# \{ x \in \mathcal{O}_F : 0 < |x^\sigma| \leq A_j \} \ll A.
\]

Indeed, (4.1) implies that for each principal ideal \( (x) \) of norm \( \text{nr}_F x \leq A \) there are \( O(\log(2 + A/\text{nr}_F x)^{d-1}) \) generators satisfying the size constraints in (4.2), hence the left hand side of (4.2) is at most

\[
\ll \sum_{\text{nr}_F(a) \leq A} \log \left( 2 + \frac{A}{\text{nr}_F(a)} \right)^{d-1} \ll A.
\]

We remark that the estimate (4.2) is a trivial lattice point count if all \( A_j \gg 1 \). It is a little less trivial if some \( A_j \) are very large and others are very small.

We use the notation \( r_Q(\ell) \) to denote the number of integral representations of \( \ell \) by \( Q \). From now we consider quadratic forms in 2, 3 and 4 variables. The following lemma derives uniform estimates for representation numbers of quadratic forms, averaged over thick and not so thick sequences (the thinner sequences are needed because of the special form of our amplifier). The lemma will be used in Section 5.2 when we derive the volume bound.

**Lemma 2.** Let \( Q \) be a totally positive-definite integral quaternary quadratic form of determinant \( \Delta \) and level \( \mathfrak{n} \). Let \( y, y_1, y_2 > 1 \). Then

(4.3)

\[
\sum_{\ell \in \mathcal{O}_F} r_Q(\ell) \ll \frac{y^2}{\text{nr}_F(\Delta)^{1/2}} + \frac{y^{3/2}}{\text{nr}_F(\Delta)/\text{nr}_F(\mathfrak{n})^{1/2}} + y.
\]

If in addition \( h_1 \asymp 1 \) in (3.2), then

(4.4)

\[
\sum_{\ell_1 \in \mathcal{O}_F} \sum_{\ell_2 \in \mathcal{O}_F} r_Q(\ell_1 \ell_2) \ll y_1 \left( \frac{(y_1 y_2^2)^{3/2}}{\text{nr}_F(\Delta)^{1/2}} + \frac{y_1}{\text{nr}_F(\Delta)/\text{nr}_F(\mathfrak{n})^{1/2}} + (y_1 y_2^{3/2})^{1/2} \right) (y_1 y_2 \text{nr}_F(\Delta))^{\varepsilon},
\]

(4.5)

\[
\sum_{\ell \in \mathcal{O}_F} r_Q(\ell^2) \ll \left( \frac{y^3}{\text{nr}_F(\Delta)^{1/2}} + \frac{y^3}{\text{nr}_F(\Delta)/\text{nr}_F(\mathfrak{n})^{1/2}} + y \right) (y \text{nr}_F(\Delta))^{\varepsilon}
\]

for any \( \varepsilon > 0 \), the implied constants depending on \( \varepsilon \) alone. Here and in the following a summation condition of the type \( 0 \leq \ell^\sigma \leq y^{1/d} \) is understood to hold for all embeddings \( \sigma \).

The first and last term on the right hand side of (4.3) and (4.5) are certainly optimal. Maybe the middle term can be improved slightly. The bound (4.4) is not best possible in general, but sufficient for our purposes.
Proof. All of these bounds are proved in a similar way. We start with (4.3). We use the representation (3.2) together with the bounds (3.4). Let \( \tilde{x}_j := x_j + \sum_{i>j} c_{ji} x_i \). By (4.2) we have \( \ll (y/n_{F}(h_j))^{1/2} \) non-zero choices for \( \tilde{x}_j \). Hence we have \( \ll (y/n_{F}(h_j))^{1/2} + 1 \) choices in total for \( \tilde{x}_j \), and hence \( \ll (y/n_{F}(h_j))^{1/2} + 1 \) choices for \( x_j \) getting a bound

\[
\ll \left( \frac{y}{n_{F}(h_4)} + 1 \right)^{1/2} \left( \frac{y}{n_{F}(h_3)} + 1 \right)^{1/2} \left( \frac{y}{n_{F}(h_2)} + 1 \right)^{1/2} \left( \frac{y}{n_{F}(h_1)} + 1 \right)^{1/2}
\]

\[
\ll \frac{y^2}{n_{F}(\Delta)^{1/2}} + \frac{y^{3/2}}{(n_{F}(\Delta)/n_{F}(n))^{1/2}} + y
\]

by (3.4), (3.5) and (3.7).

In order to prove (4.5), we choose as before \( x_4, x_3, x_2 \) in

\[
\ll \left( \frac{y^2}{n_{F}(h_4)} + 1 \right)^{1/2} \left( \frac{y^2}{n_{F}(h_3)} + 1 \right)^{1/2} \left( \frac{y^2}{n_{F}(h_2)} + 1 \right)^{1/2} \ll \frac{y^3}{n_{F}(\Delta)^{1/2}} + \frac{y^2}{(n_{F}(\Delta)/n_{F}(n))^{1/2}} + y
\]

ways. Here we used again (3.4), (3.5) and (3.7). Note that \( Q(x) = \ell^2 \) implies \( x_j^2 \ll y^{1/d} \) for all \( \sigma \), since \( \lambda_{l}^2 \gg 1 \) by (3.6). Once we have fixed \( x_2, x_3, x_4 \) we are left with counting pairs \((x_1, \ell)\) satisfying

\[
2a_{11} \ell^2 - (a_{11} x_1 + \xi)^2 = D
\]

where

\[\xi = \sum_{j=2}^{4} a_{1j} x_j, \quad D = - \sum_{i,j=2}^{4} (a_{1i} a_{1j} - a_{11} a_{ij}) x_i x_j.\]

Note that \( \xi \ll y^{1/d} \) by (3.3) and our assumption \( h_1 \gg 1 \), and \( D^\sigma \ll (\Delta^\sigma)y^{2/d} \) by the same argument. It follows that \( x_j^2 \ll (\Delta^\sigma)^{1/2} y^{1/d} \) (although precise exponent are insignificant here – we only need polynomial dependence).

Let us first assume that \( 2a_{11} = b^2 \), say, is a square in \( F \) and hence in \( \mathcal{O}_F \). If \( D \neq 0 \), then by a standard divisor argument there are \( \ll (n_{F}D)^\varepsilon \) pairs of principal ideals \((bk - a_{11} x_1 - \xi,\quad (bk + a_{11} x_1 + \xi)\) whose product equals \( (D) \), and by (4.1) each of these has \( \ll (y n_{F}(\Delta))^\varepsilon \) generators \( g \in \mathcal{O}_F \) satisfying \( g^\sigma \ll (\Delta^\sigma)^2 y^{2/d} \). This in turn gives \( \ll (y n_{F}(\Delta))^\varepsilon \) choices for \( x_1 \).

If \( D = 0 \), we choose \( \ell \) freely in \( \mathcal{O}(y) \) ways (by (4.2)), and then there are at most two choices for \( x_1 \). We determine how often the case \( D = 0 \) happens. The quantity \( D \) is a definite ternary quadratic form in \( x_2, x_3, x_4 \) whose determinants of its upper left \( k \times k \) submatrices \( (1 \leq k \leq 3) \) are precisely the determinants of \((k+1) \times (k+1)\) upper left submatrices of \( A \) (up to sign). In particular we see that \( D = 0 \) if and only if \( x_2 = x_3 = x_4 = 0 \).

Let us now assume that \( 2a_{11} \) is not a square in \( F \). Then \( D \neq 0 \), and we need to solve a Pell-type equation. There are \( \ll (n_{F}D)^\varepsilon \) ideals \((\sqrt{2a_{11}} - a_{11} x_1 - \xi,\quad \sqrt{2a_{11}} + a_{11} x_1 + \xi)\) in the totally real field \( E = F(\sqrt{2a_{11}}) \) of relative norm \( D \), and again by (4.1) each of these yield \( \ll (y n_{F}(\Delta))^\varepsilon \) solutions for \( x_1 \). This establishes (4.5).

Finally we prove (4.4). Again we fix \( x_1, x_3, x_2 \) as above, and we fix \( \ell_1 \). This gives a total count of

\[
y_1 \left( \frac{(y_1 y_2)^{3/2}}{(n_{F}(\Delta)^{1/2}} + \frac{y_1 y_2}{(n_{F}(\Delta)/n_{F}(n))^{1/2}} + \frac{(y_1 y_2)^{1/2}}{} \right),
\]

and we are left with counting pairs \((x_1, \ell_2)\) satisfying \( 2a_{11} \ell_1 \ell_2^2 - (a_{11} x_1 + \xi)^2 = D \) with \( \xi \) and \( D \) as above. Now we argue exactly as in the previous case.

For a polynomial

\[
P(x_1, \ldots, x_k) = \sum_{n \in \mathbb{N}_{\geq 0}^k} a_n x^n \in \mathcal{O}_F[x_1, \ldots, x_k]
\]
let $H(P) := \sum_1^n|a_n^P|$ denote its “height”. For $\ell \in \mathcal{O}_F$ we write $|\ell|_\infty := \max_\sigma|\ell^\sigma|$. The next two lemmas have preparatory character and are used as in input for the important Lemma 5 below. However, in particular Lemma 3 may be of independent interest, as it bounds uniformly representation numbers of positive definite ternary and quaternary forms over (fixed) totally real number fields.

**Lemma 3.** a) Let $P(x,y) \in \mathcal{O}_F[x,y]$ be a quadratic polynomial and assume that its quadratic homogeneous part is a totally positive definite quadratic form. Let $\ell \in \mathcal{O}_F$. Then there are $\ll (H(P)(1 + |\ell|_\infty)^r)$ solutions to $P(x,y) = \ell$.

b) Let $Q$ be a totally positive definite integral ternary quadratic form over $F$ of discriminant $\Delta$ and let $\ell \in \mathcal{O}_F \setminus \{0\}$. Then $r_Q(\ell) \ll \nu_F(\ell)^{1/2}(|\ell|_\infty \nu_F(\Delta))^{r}.$

c) Let $Q$ be a totally positive definite integral quaternary quadratic form over $F$ of discriminant $\Delta$ and let $\ell \in \mathcal{O}_F \setminus \{0\}$. Then $r_Q(\ell) \ll \nu_F(\ell) (|\ell|_\infty \nu_F(\Delta))^{r}.$

Here all implied constants depend on $\varepsilon$ at most.

**Remark 4.1.** The proof gives slightly stronger bounds for parts b) and c); for instance in the situation of part b) we obtain $r_Q(\ell) \ll (\nu_F(\ell)^{1/3} \nu_F(\Delta)^{-1/6} + |\ell|_\infty \nu_F(\Delta)^{r})$, but we do not need these refinements. We use part a) of the lemma in parts b) and c). Although they could also be proved without recourse to the first part, the first part will be needed in the proof of Lemma 5, and so we take the opportunity to state it and prove it here.

**Proof.** a) We write $P(x,y) = Q(x,y) + L(x,y) + C = \ell$ where $Q(x,y) = ax^2 + bxy + cy^2$ with $a \neq 0$ is a totally positive quadratic form over $\mathcal{O}_F$ of discriminant $\Delta$, $L(x,y) = \alpha_3 x + \alpha_2 y$ is a linear form over $\mathcal{O}_F$, and $C \in \mathcal{O}_F$. Let $\xi = (\beta - 2\alpha)/\Delta$, $\eta = (b\alpha - 2a\beta)/\Delta$. A little high school algebra shows

$$P(x,y) = \left(2a(x + \xi) + b(y + \eta)\right)^2 - \Delta(y + \eta)^2 + P(-\xi, -\eta).$$

Hence by a linear change of variables the equation $P(x,y) = \ell$ is equivalent to

$$X^2 - \Delta Y^2 = 4a\Delta^2(\ell - P(-\xi, -\eta)),$$

where $X$ and $Y$ satisfy certain congruence conditions modulo $4|\Delta|^2$. Clearly the norm of the left hand side is polynomial in $H(P)$ and $|\ell|_\infty$. We are now left with a norm form equation of the totally imaginary field $E = F(\sqrt{-|\Delta|})$ over the totally real field $F$, and the result follows.

b) We use the representation (3.2) with $n = 3$ together with the bounds (3.3), (3.4), (3.5). By (4.2) we can choose $x_3$ in $\ll (\nu_F(\ell)/\nu_F(\Delta^{1/3}) + 1)^{1/2} \ll (\nu_F(\ell)/\nu_F(\Delta)^{1/3} + 1)^{1/2}$ ways, and are left with an inhomogeneous binary problem for which part (a) applies.

c) This is proved in the same way. We choose $x_4$ and $x_3$ and are left with a binary problem. $\square$

In the following lemma we denote by $\|\|_2$ the usual Euclidean norm on $\mathbb{R}^n$, which is (in general) not induced by the inner product (3.1).

**Lemma 4.** Let $Q(x) = \frac{1}{2}x^tAx$ be a positive definite ternary quadratic form with real coefficients and eigenvalues $\gg 1$, let $x \in \mathbb{R}^3$ be such that $Q(x) = 1$ and let $\ell > 0$, $\eta > 0$.

a) If $Q(y) = 1$ and $\langle y, x \rangle^2 \geq 1 - \eta$, then $\min_{x} \|y \pm x\|_2 \ll \eta^{1/2}$.

b) Let $Q(y_i) = \ell$ for $i = 1, 2, 3$ and assume $|\langle y_i, x \rangle| \leq \ell^{1/2} \eta^{1/2}$. Then $\det(y_1, y_2, y_3) \ll \ell^{3/2} \eta^{1/2}$.

c) Let $Q(y) = \ell$, then $\|y\|_2 \ll \ell^{1/2}$.

All implied constants are absolute (if the underlying number field is fixed).

**Proof.** The assertions are clear if $Q(y) = y_1^2 + y_2^3 + y_3^2$ and $x = (0, 0, 1)^t$ is the north pole. In the general case, we write $\frac{1}{2}A = B^tB$ for some unique positive symmetric matrix $B \in \text{GL}_3(\mathbb{R})$, so that $\|Bx\|_2 = 1$. Let $S \in \text{O}_3(\mathbb{R})$ be any orthogonal matrix with $SBx = (0, 0, 1)^t$. Since $A$ has eigenvalues $\gg 1$, the same holds for $B$ and hence for $SB$. For the proof of a) in the general case we conclude $\|y - x\|_2 \ll \|SB(y - x)\|_2$, and for the two vectors $SBx$, $SBx$ the above special case applies. The
other two parts are proved in the same way.

The rather complicated proof of the next lemma follows to some extent the argument in [Van97, Lemma 2.1]. It will be used for the bound in the eigenvalue aspect in Section 5.3. Here and in the following we extend the norm \( n_F \) in an obvious way to a function \( \mathbb{R}^d \to \mathbb{R}_{\geq 0} \).

**Lemma 5.** Let \( Q(y) = y_0^2 + \tilde{Q}(\tilde{y}) \) with \( \tilde{y} = (y_1, y_2, y_3) \) be an integral positive definite quaternary quadratic form over \( \mathcal{O}_F \) of discriminant \( \Delta \), and let \( \ell \in \mathcal{O}_F \setminus \{0\} \) be totally positive. Assume that \( \tilde{Q} \) is reduced in the sense of Section 3. Let \( \eta = (\eta_1, \ldots, \eta_d) \in (0,1]^d \). Let \( x_1, \ldots, x_d \in \mathbb{R}^3 \) satisfy \( \tilde{Q}^\sigma_j(x_j) = 1 \). Then the following two estimates hold:

\[
\# \{ y \in \mathcal{O}_F^3 \mid Q(y) = \ell, \langle y_0^2, x_j \rangle^2 \leq \eta_j \ell^\sigma_j \} \ll \left( n_F(\eta)^{1/2} n_F(\ell) + 1 + \min \left( n_F(\ell)^{3/2} n_F(\eta)^{1/2}, n_F(\ell)^{1/2} \right) \right) (|\ell|^{1/2} n_F(\Delta))^{\varepsilon}. \tag{4.6}
\]

and

\[
\# \left\{ y \in \mathcal{O}_F^3 \mid Q(y) = \ell, \tilde{Q}^\sigma_j(\tilde{y}^\sigma_j) - \langle \tilde{y}^\sigma_j, x_j \rangle^2 \leq \eta_j \ell^\sigma_j \right\} \ll \left( 1 + \min \left( n_F(\ell)^{3/11} n_F(\eta)^{12/11}, n_F(\ell)^{1/11} \right) \right) (|\ell|^{1/2} n_F(\Delta))^{\varepsilon}. \tag{4.7}
\]

As before, all implied constant depend on \( \varepsilon \) at most.

**Proof.** We will frequently use Lemma 4 which is applicable because of (3.6).

We start with the proof of (4.6). By (4.2) there are \( \ll (n_F(\eta) n_F(\ell))^{1/2} \) choices for \( y_0 \neq 0 \), and for each of them there are by Lemma 3b at most \( \ll n_F(\ell)^{1/2} (|\ell|^{1/2} n_F(\Delta))^{\varepsilon} \) choices for \( \tilde{y} \). This gives the first term in (4.6).

We proceed to count the solutions with \( y_0 = 0 \). There are at most 2 linearly dependent solutions to \( \tilde{Q}(\tilde{y}) = \ell \) (namely \( \tilde{y} \) and \( -\tilde{y} \)), hence after adding 1 to the count of (4.6) we can assume that there are at least two linearly independent solutions \( \tilde{y}_1 = (y_{11}, y_{12}, y_{13})^t \), \( \tilde{y}_2 = (y_{21}, y_{22}, y_{23})^t \), say, satisfying

\[
\tilde{Q}(\tilde{y}_j) = \ell, \quad \langle \tilde{y}_j, x_j \rangle^2 \leq (\eta_j \ell^\sigma_j)^2, \quad j = 1, 2, \ldots, d.
\]

Recall that by Lemma 4c any solution \( y \) to (4.8) satisfies \( y^2 \ll (\ell^\sigma)^{1/2} \) for \( 1 \leq i \leq 3 \) and all \( \sigma \). Now any other solution \( \tilde{y}_3 \) satisfies \( \det(\tilde{y}_j, \tilde{y}_3^\sigma) \ll (\ell^\sigma)^{3/2} \eta_j^{1/2} \) by Lemma 4b, as well as \( \tilde{Q}(\tilde{y}_3) = \ell \).

By (4.2) there are \( 1 + n_F(\ell)^{3/2} n_F(\eta)^{1/2} \) choices for the determinant (including 0). For a fixed value of the determinant and some \( \tilde{y}_3 \in \mathcal{O}_F^3 \) yielding this value, all vectors yielding this determinant are of the form \( \tilde{y}_3 = \tilde{b} + a\tilde{y}_1 + b\tilde{y}_2 \in \mathcal{O}_F^3 \) with \( a, b \in K \). Let \( \mathbf{Y} := (Y_{12}, Y_{13}, Y_{12}) := \tilde{y}_1 \times \tilde{y}_2 \neq 0 \) be the cross product which is non-zero, since \( \tilde{y}_1 \) and \( \tilde{y}_2 \) are linearly independent. Assume without loss of generality that \( Y_{12} = y_{11} y_{22} - y_{12} y_{21} \neq 0 \), the other cases being analogous. Then \( a\tilde{y}_1 + b\tilde{y}_2 \in \mathcal{O}_F^3 \) implies that both \( \alpha := Y_{12} a \) and \( \beta := Y_{12} b \) are in \( \mathcal{O}_F \), and \( \tilde{Q}(Y_{12} \tilde{y}_3) = \tilde{Q}(Y_{12} \tilde{y}_3 + \alpha \tilde{y}_1 + \beta \tilde{y}_2) = Y_{12}^2 \ell \) is a homogeneous binary problem in \( \alpha, \beta \) with polynomial height in \( |\ell| \) and the coefficients of \( Q \) (recall that \( Q \) is a reduced form), hence there are \( \ll (|\ell|^{1/2} n_F(\Delta))^{\varepsilon} \) solutions by Lemma 3a.

Alternatively, by Lemma 3b we have the trivial bound \( \ll n_F(\ell)^{1/2} (|\ell|^{1/2} n_F(\Delta))^{\varepsilon} \) for the number of solutions with \( y_0 = 0 \). Combining these two counts gives the last term in (4.6).

We proceed to prove (4.7). Let \( \delta_{1,j} \delta_{2,j} = \eta_j \). We will fix \( \delta_{1,j}, \delta_{2,j} \) later and assume for the moment only \( n_F(\delta_1), n_F(\delta_2) \ll n_F(\ell) \). It is enough to estimate the number of \( y \in \mathcal{O}_F^3 \) with \( Q(y) = \ell \) satisfying \( \tilde{Q}^\sigma_j(\tilde{y}^\sigma_j) \leq \delta_{1,j} \ell^\sigma_j \) and the number of \( y \in \mathcal{O}_F^3 \) with \( Q(y) = \ell \) satisfying

\[
(1 - \langle \tilde{y}^\sigma_j, x_j \rangle^2 / \tilde{Q}^\sigma_j(\tilde{y}^\sigma_j)) \leq \delta_{2,j}.
\]
\(y = (y_0, \tilde{y}), z = (z_0, \tilde{z}) \in \mathcal{O}_F^1\) with \(Q(y) = Q(z) = \ell\) equivalent if \(\tilde{z} = c \tilde{y}\) for some \(c \in K\). We claim that the cardinality of each equivalence class \([y]\) is small. This can be seen as follows: Clearly \(c) \subseteq (y_1, y_2, y_3)\). Fix a fractional ideal \(a \supseteq \mathcal{O}_F\) with \(\text{nr}_F(a) \ll 1\) in the ideal class of \((y_1, y_2, y_3)\). Then \(c) \subseteq (y_1, y_2, y_3)\) at \(a = (\alpha)\), say, where \(\text{nr}_F(a) \gg \text{nr}_F(\ell)^{-1/2}\) and we choose a generator such that \(|\alpha|^2 \gg \text{nr}_F(a)^{1/4}\), say. Hence we can write \(c = da\) with \(d \in \mathcal{O}_F\). After multiplication with \(1/\alpha\), the equation \(Q(z) = \ell\) becomes an integral binary problem in \(d\) and \(z_0\) which by Lemma 3a has \(\ll (|\ell| \text{nr}_F(\Delta))^{\varepsilon}\) solutions. It is therefore enough to count the number of equivalence classes, and to this end we pick a set of representatives \(y\); then by construction the corresponding vectors \(\tilde{y}\) are pairwise not collinear.

From these representatives we select a vector \(y \in \mathcal{O}_F^1\) such that
\[
Y := \prod_{\sigma} \max(|(y_1)_{\sigma}|, |(y_2)_{\sigma}|)
\]
is maximal. We may assume \(Y \neq 0\). By Lemma 4c we have \(Y \ll \text{nr}_F(\ell)^{1/2}\). For any other (non-equivalent) solution \(z = (z_0, \ldots, z_3) = (z_0, \tilde{z}) \in \mathcal{O}_F^1\), we conclude from (4.9) that
\[
\left\langle \frac{\tilde{y}}{Q_{\sigma_j}(\tilde{y})^{1/2}}, \frac{\tilde{z}}{Q_{\sigma_j}(\tilde{z})^{1/2}} \right\rangle \gg 1 - \delta_{2,j},
\]
and hence by Lemma 4c
\[
|(y_1z_2 - y_2z_1)^{\sigma_j}| \leq \|\tilde{y}\|_2 \min_{\pm} \|\tilde{y} \pm \tilde{z}\|_2 \leq \|\tilde{y}\|_2 \|\tilde{z}\|_2 \left(\frac{\text{nr}_F(\ell)}{\text{nr}_F(1/2)}\right)^{1/2} \ll \delta_{2,j}^{1/2} \|\tilde{y}\|_2 \|\tilde{z}\|_2.
\]
We first count solutions \(z\) where \(y_1z_2 - y_2z_1 \neq 0\). Then the non-zero principal ideal \((y_1z_2 - y_2z_1)\) has norm
\[
\ll D := \text{nr}_F(\delta_2)^{1/2} \prod_{\sigma} \|\tilde{y}\|_2 \|\tilde{z}\|_2 \ll \text{nr}_F(\delta_2)^{1/2} \text{nr}_F(\ell)^{1/2}
\]
and it is divisible by the ideal \(\eta := (y_1, y_2)\). There are \(\ll D/\text{nr}_F(\eta)\) such ideals. For each of these ideals the number of generators satisfying (4.10) is \(\ll \text{nr}_F(\ell)\) by (4.1). Let us fix one such generator \(g\) and count the number of \(z_1, z_2\) satisfying \(y_1z_2 - y_2z_1 = g\) and \(z_1^* / z_2^* \ll (\ell) / (2)\) (by Lemma 4c). Assume there is a solution \((z_1, z_2)\). If \((z_1, z_2)\) is another solution, write \((z_1^*, z_2^*) := (z_1 - z_1, z_2 - z_2)\). Then we need to count the number of \((z_1^*, z_2^*) \neq (0, 0)\) satisfying \((z_1^* / z_2^*) \ll (\ell) / (2)\) and \(y_1z_2^* = y_2z_1^*\). Fix an integral ideal \(a\) in the ideal class of \(\eta^{-1}\), and write \(\eta a = (\alpha)\) with \(\text{nr}_F(a) \gg \text{nr}_F(\eta)\). Then the fractional ideal \((y_2 / \alpha)\) divides \((z_2^*)\), and hence \(z_2^* = y_2w_2 / \alpha\) for some \(w_2 \in \mathcal{O}_F\). Similarly \(z_1^* = y_1w_1 / \alpha\) for some \(w_1 \in \mathcal{O}_F\). Inserting into the equation \(y_1z_2^* = y_2z_1^*\) we see \(w_1 = w_2 = w \neq 0\) say, and we have the bound \(w \ll (\ell) / (2)\) max \(\text{nr}_F(\ell)^{1/2}\) by (4.2). The number of such \(w\) is \(\ll \text{nr}_F(\ell)^{1/2} \text{nr}_F(\eta) \ll \text{nr}_F(\ell)^{1/2} \text{nr}_F(\eta)^{-1}\). Thus the total number of solutions \((z_1, z_2)\) with \(y_1z_2 - y_2z_1 \neq 0\) is
\[
\ll \text{nr}_F(\ell)^{3/2} \text{nr}_F(\delta_1)^{1/2} \text{nr}_F(\ell)^{-1}
\]
Alternatively, by the definition of \(\tilde{y}\) and \(Y\), the principal ideals \((z_1)\) and \((z_2)\) have norm at most \(Y\), and the number of generators whose conjugates are bounded by \(O(\ell)\) is by (4.1) at most \(\ll \text{nr}_F(\ell)^{1/2}\). Hence the number of \((z_1, z_2)\) is at most \(\ll \text{nr}_F(\ell)^{1/2}\). Together with (4.11) we get the bound
\[
\ll \text{nr}_F(\ell)^{1/2} \text{nr}_F(\delta_1)^{1/3}
\]
Once \(z_1, z_2\) are fixed, the equation \(Q(z) = \ell\) is a binary problem in \((z_0, z_3)\), thus by Lemma 3a our total count of (equivalence classes) \(z \in \mathcal{O}_F^1\) with \(y_1z_2 - y_2z_1 \neq 0\) is
\[
\ll \text{nr}_F(\delta_2)^{1/3} \text{nr}_F(\ell)^{1/2} \ll \text{nr}_F(\delta_2)^{1/3} \text{nr}_F(\ell)^{1/2}
\]
We now consider the set of all those (equivalence classes of) solutions \( z \in O_F^4 \) with \( y_1z_2 - y_2z_1 = 0 \). For any two such solutions \( z, w \) we have by construction that \( z, w \) are not collinear, hence \( z_1z_3 - z_3z_1 \neq 0 \). From the set of these vectors \( z \) we select some \( z^* \) such that

\[
Z := \prod_{\sigma} \max((z_1^*)^\sigma, |(z_3^*)^\sigma|)
\]

is maximal. An argument identical to the above shows that the number of the (equivalence classes) \( z \in O_F^4 \) such that \( y_1z_3 - y_2z_1 = 0 \) satisfies the bound (4.12).

It remains to count the number of \( y \in O_F^4 \) with \( Q(y) = \ell \) satisfying \( \hat{Q}(\hat{y}^\sigma) \leq \delta_{1,\ell} \ell^\varepsilon \). By Lemma 3b the number of \( \hat{y} \) is

\[
(4.13) \quad \ll 1 + (\text{nr}_F(\delta_1)\text{nr}_F(\ell))^{3/2}(\max_j |\delta_{1,j} \ell^\varepsilon| \text{nr}_F(\Delta))^{\varepsilon},
\]

and then \( y_0 \) is determined up to sign. Choosing \( \delta_{1,j} = \eta_4^{j/11}(\ell^\varepsilon)^{-3/11} \), the sum of (4.12) and (4.13) gives the first bound of (4.7). The second bound follows trivially from Lemma 3c. \( \square \)

5. Application of the pre-trace formula

5.1. The general set-up. As in [BM11], the proof of Theorem 3 (which implies Theorem 1) follows from an application of a pre-trace formula which we recall now. Our aim in this section is to state the pre-trace formula and to construct the amplifier. This will finally result in the important bound (5.1) to which we can apply the results from the preceding section. The pre-trace formula is the spectral expansion of an automorphic kernel and works verbatim as in [BM11, Section 5.1] where the case \( F = \mathbb{Q} \) is considered and to which we refer for more details. It features the matrix coefficients \( p_{m,1} \) defined in Section 2.3. Similar non-adelic treatments can be found in [IS95, Van97]. The (only) new input compared to [BM11] is that we choose an amplifier that has support on “balanced” algebraic integers (i.e. algebraic integers whose various conjugates have roughly the same size) that generate principal prime ideals.

Fix weights \( k \in \mathbb{Z}^d \) and a Hecke character \( \chi \) as in Theorem 3 (in particular \( \chi \) is unramified at every place not dividing \( \text{disc}(O) \)). For \( \alpha = (\alpha_v)_v \in \mathcal{B}(A_f) \) such that \( \alpha_v \in O_v^\times \) for \( v \mid \text{disc}(\mathcal{O}) \), let \( f_\alpha \) be the function supported on \( Z(A_f) \hat{\mathcal{O}}^\times \alpha \hat{\mathcal{O}}^\times \) that is 1 on \( \hat{\mathcal{O}}^\times \alpha \hat{\mathcal{O}}^\times \) and satisfies

\[
f_\alpha(\lambda h) = \pi(\lambda)f_\alpha(h), \quad \lambda \in Z(A_f).
\]

Let \( \mathcal{H}(\mathcal{O}) \) be the convolution\(^3\) (spherical) Hecke algebra generated by these bi-\( \hat{\mathcal{O}}^\times \)-invariant functions \( f_\alpha \) on \( \mathcal{B}^\times(A_f) \). This algebra is commutative, and it follows from our assumptions (by \( \hat{\mathcal{O}}^\times \)-invariance) that \( \varphi \) is an eigenfunction for the action of \( \mathcal{H}(\mathcal{O}) \) by convolution: for \( f \in \mathcal{H}(\mathcal{O}) \) one has

\[
f \ast \varphi(g) = \int_{\mathcal{P}B^\times(A_f)} f(h)\varphi(gh) \, dh = \lambda_\varphi(f)\varphi(g).
\]

Given such an \( f \), we consider more generally the convolution operator on \( L^2(\mathcal{B}^\times(F)\backslash\mathcal{B}^\times(A_f), \chi) \)

\[
R(f): \psi(g) \mapsto R(f)(\psi)(g) = \int_{\mathcal{P}B^\times(A_f)} f(h)f p_{m,1}(h) \psi(gh) \, dh
\]

where we use as before the notation \( h = h_f h_\infty \in \mathcal{P}B^\times(A_f) \) and \( p_{m,1} \) was defined in (2.6). This is an integral operator with kernel given by

\[
K_f(g, g') = \sum_{\gamma \in \mathcal{P}B^\times(F)} f(g^{-1}g_f^j p_{m,1}(g^{-1}g_f^j g_\infty)).
\]

\( ^3\)for the convolution \( f_1 \ast f_2(g) = \int_{\mathcal{P}B^\times(A_f)} f_1(h)f_2(h^{-1}g) \, dh \).
It decomposes into an orthonormal (finite) basis \( \{ \psi \} \) of \( \hat{\Theta}^x K_\infty \)-invariant \( \mathcal{H}(\mathcal{O}) \)-eigenfunctions of weight 1 with respect to \( K_\infty \) containing \( \varphi \), and from the normalization of Haar measures in Section 2.4.3 one finds that

\[
K_f(g, g') = \frac{1}{d_m} \sum_{\psi} \lambda_\psi(f) \psi(g) \overline{\psi(g')}.
\]

Choosing \( f \) appropriately, one can assume that \( \lambda_\psi(f) \geq 0 \) for any such \( \psi \) and that \( \lambda_\varphi(f) \) is positive (and large): this is the principle of the amplification method. Taking \( g = g' \), we obtain

\[
|\varphi(g)|^2 \leq \frac{|m|}{\lambda_\varphi(f)} \sum_{\gamma \in \text{PB}^x(F)} |f(g_f^{-1} \gamma g_f)| p_m, 1(g_{\infty}^{-1} \gamma g_\infty)|.
\]

(Recall that \( |m| \geq 1 \) even for \( m = 0 \), which lightens our notation.)

We now construct the amplifier \( \lambda_\varphi(f) \) by a slight generalization of [BM11, §5.2] to the number field \( F \) which takes into account that the group of units \( \mathcal{O}_F \) is in general infinite. Let \( F^\times_{\infty,+} = \{ (x_\sigma)_{\sigma | \infty} \in F^\times_{\infty}, x_\sigma > 0 \} \) be the identity component of \( F^\times_{\infty} \). We fix (once and for all) a fundamental domain \( D_0 \) for the action of the totally positive units \( U^+ \) on the hyperboloid \( \{ y \in F^\times_{\infty,+} \} \).

Let \( F^\text{diag}_{\infty,+} = \{ (y, \ldots, y) \in F^\times_{\infty,+} \} \). Then the cone \( \mathcal{D} := F^\text{diag}_{\infty,+} D_0 \) is a fundamental domain for the action of \( U^+ \) on \( F^\times_{\infty,+} \).

Given some parameter \( L \geq 1 \), consider the four sets

\[
\mathcal{L}_1 := \{ \ell \in \mathcal{O}_F \cap \mathcal{D} | \text{nr}_F(\ell) \in [L, 2L] \},
\]

\[
\mathcal{L}_2 := \{ \ell \in \mathcal{O}_F \cap \mathcal{D} | \text{nr}_F(\ell) \in [L^2, (2L)^2] \},
\]

\[
\mathcal{L}_3 := \{ \ell_1 \ell_2 L \leq \text{nr}_F(\ell_1), \text{nr}_F(\ell_2) \leq 2L, \ell_1, \ell_2 \in \mathcal{O}_F \cap \mathcal{D} \},
\]

\[
\mathcal{L}_4 := \{ \ell^2 | L^2 \leq \text{nr}_F(\ell) \leq (2L)^2, \ell \in \mathcal{O}_F \cap \mathcal{D} \},
\]

where \( \ell, \ell_1, \ell_2 \) denote generators contained in \( \mathcal{D} \) of integral principal prime ideals \( p \subseteq \mathcal{O}_F \) co-prime with \( \text{disc}(\mathcal{O}) \). For any \( \ell \in \mathcal{L}_i \), let \( \alpha(\ell) = (\alpha(\ell)_v)_{\nu} \in B^x(K_F) \) be such that \( \alpha(\ell)_v = (\ell_0^{1}) \in \mathcal{O}_{F,v}^x \) if \( v \nmid \text{disc}(\mathcal{O}) \) and \( \alpha(\ell)_v = 1 \) if \( v | \text{disc}(\mathcal{O}) \). Our selected function \( f \in \mathcal{H}(\mathcal{O}) \) is then the same linear combination of the \( f_\alpha(\ell) \) as in [BM11, p. 25]. Precisely, for \( r \in \mathcal{O}_F \) let

\[
c_r = \begin{cases} 
\text{sgn}(\lambda_\varphi(f_\alpha(\ell))), & r = \ell \in \mathcal{L}_1, \\
\text{sgn}(\lambda_\varphi(f_\alpha(\ell))), & r = \ell^2, \ell \in \mathcal{L}_1, \\
0, & \text{otherwise},
\end{cases}
\]

let

\[
\hat{f} = \sum_{\ell} c_\ell f_\alpha(\ell)
\]

and define

\[
f = \hat{f} \ast \hat{f} = \sum_{d, \ell_1, \ell_2} c_{d, \ell_1, \ell_2} f_\alpha(\ell_1, \ell_2).
\]

For \( g \in \text{PB}^x(\mathbb{A}) \) let \( \mathcal{O}_g \) be the order

\[
\mathcal{O}_g := B(F) \cap g_f^{-1} \hat{\Theta} g_f
\]

(where as before \( g_f \) is the finite adèle, the index does not refer to the function \( f \)). Observe that the \( \gamma \in \text{PB}^x(F) \) satisfying \( f_\alpha(\ell)(g_f^{-1} \gamma g_f) \neq 0 \) correspond to \( \gamma \in \mathcal{O}_g \) with \( \text{nr}(\gamma) = \ell \). This remark along with our choice of amplifier yields the bound

\[
|\varphi(g)|^2 \ll |m| L^2 \left( \frac{1}{L^2} + \sum_{i=1}^{4} \frac{1}{L^{2+1/2}} \sum_{\ell \in \mathcal{L}_i} \sum_{\gamma \in \mathcal{O}_g, \text{nr}(\gamma) = \ell} |p_{m, 1}(g_{\infty}^{-1} \gamma g_\infty)| \right).
\]

Before we proceed, we make the important observation that the quadratic form \( Q_g \) associated to the order \( \mathcal{O}_g \) is, in a suitable basis, of the form \( y_0^2 + \) a ternary form in \( y_1, y_2, y_3 \), the latter
corresponding to the restriction of the norm form to the traceless quaternions. In particular, the assumptions of Lemma 2 and Lemma 5 are satisfied. Moreover, by the discussion in Section 2.1, we know that \( \mathcal{O}_g \) is everywhere locally conjugate to the order \( \mathcal{O} \), and hence \( \mathcal{O}_g \) is an Eichler order and its associated quadratic form \( Q_g \) has the same discriminant \( \Delta \) and the same level \( n \) as \( Q = (\mathcal{O}, \text{nr}) \).

We set
\[
V := \text{nr}_F(\text{disc} \mathcal{O}_g)^{1/2} = \text{nr}_F(\Delta)^{1/2} = \text{nr}_F(n),
\]
where the last equality follows from (2.1) and (2.2), and we note that \( V = V_2^{1+o(1)} \) by (1.4).

5.2. **Bound in the volume aspect.** By the trivial bound \( |p_{m,1}(g_{\infty}^{-1}\gamma g_{\infty})| \leq 1 \), we obtain
\[
|\varphi(g)|^2 \ll |m| L^{\frac{1}{2}} \left( \frac{1}{L} + \sum_{i=1}^{4} \frac{1}{L^{2+i/2}} \sum_{\ell \in \mathcal{L}_i} r_{Q_{\ell}}(\ell) \right)
\]
where as before \( Q_g \) is the quaternary quadratic form associated with the order \( \mathcal{O}_g \). We use Lemma 2 with \( \text{nr}_F(\Delta) = V_2^2 \), \( \text{nr}_F(\Delta)/\text{nr}_F(n) = V \). More precisely, we use (4.3) for \( i = 1 \) resp. \( i = 2 \) with \( y \ll L \) resp. \( y \ll L^2 \), we use (4.4) for \( i = 3 \) with \( y_1 = y_2 \ll L \) and we use (4.5) for \( i = 4 \) with \( y \ll L^2 \). In this way we obtain
\[
|\varphi(g)|^2 \ll |m|(VL)^{\varepsilon} \left( \frac{1}{L} + \frac{L^{1/2}}{V^{1/2}} + \frac{L^2}{V} \right).
\]
Choosing \( L = V^{1/3} \) we find
\[
(5.2) \quad \varphi(g) \ll |m|^{1/2} V^{-1/6+\varepsilon}.
\]

5.3. **Bound in the eigenvalue aspect.** Let \( x' := \rho_{g_{\infty}}(x^0) \in B^0(F_{\infty}) \simeq (\mathbb{R}^3)^d \) be the \( d \)-tuple of unit vectors
\[
x'_\sigma = \rho_{g_{\infty}}(x^0) = g_{\sigma}x^0g_{\sigma}^{-1}
\]
obtained by transforming the “north pole” \( x^0_\infty \) by the rotation defined by \( g_{\infty} \). The bi-K\(_{-\infty}\)-invariance of \( |p_{m,1}(.)| \) implies that the function
\[
\gamma \mapsto |p_{m,1}(g_{\infty}^{-1}\gamma g_{\infty})|
\]
depends on \( \gamma \) (with \( \text{nr}(\gamma) = \ell \in \mathcal{O}_F \)) only through the quantity
\[
(5.3) \quad t = t(\gamma) = (t_{\sigma})_{\sigma} \in [-1, 1]^d, \quad t_{\sigma} = \langle x'_\sigma, \gamma x'_\sigma(\gamma^{-1}) \rangle_{B^0} = -1 + 2 \frac{\langle \gamma, x'_\sigma \rangle_{B}^2}{\ell_{\sigma}}.
\]
The equality can be checked by brute force computation in the real quaternion algebra \( B(\mathbb{R}) \), and by conjugation-invariance of the trace it is even enough to assume that \( x'_\sigma \) is, say, the north pole in \( B^0(\mathbb{R}) \).

By Lemma 1 we have
\[
|p_{m,1}(g_{\infty}^{-1}\gamma g_{\infty})| \ll \prod_{\sigma} \min \left( 1, m_{\sigma}^{-1/2}(1-t_{\sigma}^2)^{-1/4} \right).
\]
We cut the \( \gamma \)-sum in (5.1) into dyadic (multi-dimensional) intervals according to
\[
1 \pm \langle x'_\sigma, \gamma x'_\sigma(\gamma^{-1}) \rangle \simeq \eta_{\sigma} = 2^{-c_{\sigma}}, \quad 0 \leq c_{\sigma} < C
\]
\[
1 \pm \langle x'_\sigma, \gamma x'_\sigma(\gamma^{-1}) \rangle \leq 2^{-C}
\]
where \( 2^{-C} \simeq (LV|m|)^{-100}. \) There are \( O((LV|m|)^{\varepsilon}) \) such intervals. By (5.3) the two conditions
\[
1 \pm \langle x'_\sigma, \gamma x'_\sigma(\gamma^{-1}) \rangle \leq \eta_{\sigma} \ell_{\sigma}^2 \text{ and } (\text{the notation of} 
\]

Lemma 5) $\bar{Q}^\sigma(\bar{\gamma}^\sigma) - (\bar{\gamma}^\sigma, x^\sigma_\gamma)^2 \leq \eta_\sigma \ell^\sigma$. Here we used again the special shape of our quadratic form. Hence by Lemma 5 we have

$$1 \ll \left( 1 + \min \left( |nr_F(\ell)|^{1/2}, |nr_F(\ell)|^{3/2} |nr_F(\eta)|^{1/2} \right) + \min \left( |nr_F(\eta)|^{3/11} |nr_F(\ell)|^{12/11}, |nr_F(\ell)| \right) \right) \left( V |nr_F(\ell)|^\varepsilon \right).$$

Here we used that $|nr_F(\ell)| |nr_F(\eta)|^{1/2} \leq \min \left( |nr_F(\eta)|^{3/11} |nr_F(\ell)|^{12/11}, |nr_F(\ell)| \right)$. If $\eta_\sigma \leq 2^{-C}$ for at least one $\sigma$, then $|nr_F(\eta)| \ll (LV|m|)^{-100}$, hence those $\gamma$ contribute at most

$$|m| L^{2+\varepsilon} L^{\min(i,2)+\varepsilon} \ll \frac{|m|}{L^{1+\varepsilon}},$$

to (5.4). If $\eta_\sigma > 2^{-C}$ for all $\sigma$, then $|nr_F(\eta)| \gg (LV|m|)^{-100}$, and we are left with

$$\ll \sum_{i=1}^4 \frac{|m|}{L^{2+\varepsilon}} L^{\max(\min(i,2),0)} \max_{(LV|m|)^{-100} \leq |nr_F(\eta)| \leq 1} \left( \frac{V L |m|}{|nr_F(\eta)|} \right)^\varepsilon \min \left( 1, \frac{1}{|nr_F(\eta)|^{1/2} |nr_F(\eta)|^{1/4}} \right) \left( 1 + \min \left( L^{i/2}, L^{3i/2} |nr_F(\eta)|^{1/2} \right) + \min \left( |nr_F(\eta)|^{3/11} L^{12i/11}, L^i \right) \right).$$

We choose

$$L := |m|^{3/20}.$$

The discussion of (5.5) as a function of $\eta$ with $0 \leq |nr_F(\eta)| \leq 1$ with this choice of $L$ is elementary, but a bit tedious. The easiest way is to write $|nr_F(\eta)| = |m|^\beta$ with $\beta \leq 0$. Then (5.5) becomes

$$\ll |m| V^\varepsilon \sum_{i=1}^4 \max_{\beta \leq 0} \left( \log |m| \right) \left( 1 + \frac{3}{20} \left( \min(i,2) - \frac{i}{2} \right) + \min \left( 0, -\frac{1}{2} - \frac{1}{4} \beta \right) \right) + \max \left( 0, \min \left( \frac{i}{2}, \frac{3}{20}, \frac{3i}{2}, \frac{3}{20}, \frac{\beta}{2}, \min \left( \frac{12i}{11}, \frac{3}{20}, \frac{3\beta}{12i}, \frac{3i}{20} \right) \right) \right).$$

For each $i \in \{1, 2, 3, 4\}$, the exponent is a piecewise linear function in $\beta$ that is elementary to discuss. Figure 1 displays the four functions.

It is easy to check that the maximum is taken at $\eta \leq |m|^{-2}$ and $i = 2$ and at $\eta = |m|^{-1/5}$ and $i = 4$ in which case (5.5) is

$$\ll |m|^{17/20} (|m| V)^\varepsilon.$$

Combining (5.4) and (5.6) we obtain

$$\varphi(g) \ll |m|^{17/40} (|m| V)^\varepsilon.$$

The bounds (5.2) and (5.7) prove (1.7) (observe that $3/80 < 1/27$).
5.4. Hybrid bound. Combining (5.2) and (5.7) and writing $|m| \asymp |\lambda|^{1/2}$, we obtain
\[
\varphi(g) \ll (|\lambda|^V)^{1/4} \times (V^{-1/6} |\lambda|^{-3/80} |\lambda|^{-3/80} V^{20/29}) \\
\ll |\lambda|^{1/4} (|\lambda|^{1/2} V)^{-3/58+\varepsilon} \ll |\lambda|^{1/4} (|\lambda|^{1/2} V)^{-1/20}.
\]
This completes the proof of Theorem 3.

6. Bounds for automorphic forms on 3-dimensional ellipsoids

One of the main reasons for extending [BM11] from $\mathbb{Q}$ to general totally real number fields is that this will allow us to obtain bounds for automorphic forms associated to \textit{quaternary} quadratic spaces. As was recalled in §2.2 there is a close relationship between automorphic functions associated to \textit{quaternary} quadratic spaces and automorphic functions attached to ternary quadratic spaces but over a \textit{quadratic} extension of the base field.

In the next section we define the arithmetic manifold $X^{(3)}(\mathcal{O})$ as an adelic quotient of a special orthogonal group in four variables $G = SO(\mathcal{B}')$ on which the automorphic forms considered in Theorem 2 live. In order to bound automorphic function on $G$, we pass to automorphic forms on the somewhat larger group $\tilde{G}$, defined in (2.4). Then the scene is prepared to copy the arguments from Section 5 and derive Theorem 2.

6.1. Automorphic forms associated to orthogonal groups in four variables. From now on we restrict to the following situation: $F$ is a fixed totally real number field, $E$ a fixed totally real quadratic extension of $F$ (possibly $F \times F$), $B$ is a totally definite quaternion algebra over $F$ and $B'$ is the 4-dimensional vector space of $B \otimes_F E$ described in §2.2; as explained above, any totally definite quaternary quadratic space is similar to some $(B', \nu)$. In the following all implied constants may depend on $F$ and $E$.

In the space $(B', \nu)$ consider the following quadratic lattice: Let $\mathcal{O} \subset B$ be an Eichler order, $\mathcal{O}_E$ a maximal order in $E$ (if $E = F \times F$ we take $\mathcal{O}_E = \mathcal{O}_F \times \mathcal{O}_F$) and $\mathcal{O}_{B'E} = \mathcal{O} \otimes_F \mathcal{O}_E$; finally let $(\Lambda, \nu) = (\mathcal{O}', \nu)$ be the quadratic lattice with $\mathcal{O}' = \{ z \in \mathcal{O}_{B'E}, \sigma(z)^* = z \} \subset B'$.

It is easy to see that
\[
\text{nr}_F(\text{disc} \mathcal{O}') = \text{nr}_E(\text{disc} \mathcal{O})^{1+o_E(1)},
\]
We denote by $K_{\mathcal{O}'} \subset G(\mathbb{A}_F)$ the stabilizer of $\mathcal{O}'$ and by $K_\infty = \prod_i K_{\sigma_i} \simeq SO_3(\mathbb{R})^d$ the stabilizer of the element $1 \in B'(F_\infty)$ in $G(F_\infty) \simeq SO_4(\mathbb{R})^d$. We have the identification
\begin{equation}
X^{(3)}(\mathcal{O}) := G(F) \backslash G(\mathbb{A}) / K_{\mathcal{O}'} K_\infty \simeq \bigsqcup_{i \in \text{gen}(\mathcal{O}')} \Gamma_i \backslash (S_3)^d, \Gamma_i \subset SO_4(\mathbb{R})^d
\end{equation}
where the disjoint union is indexed by the set of classes in the genus of $(\mathcal{O}', \nu)$, and for a given representative of a genus class $(\Lambda_i, \nu)$, $\Gamma_i$ is isomorphic to the subgroup of $SO(B')(F)$ which preserves the lattice $\Lambda_i$. As above we are interested in a sup-norm bound of functions $\varphi$ on $X^{(3)}(\mathcal{O})$ which are eigenfunctions of the Laplace operator $\Delta = (\Delta_\sigma)_\sigma$ and of a suitable algebra of Hecke operators $\mathcal{H}(K_{\mathcal{O}'})$. The Laplace operator is normalized so that the eigenvalues have the shape
\begin{equation}
\lambda = (\lambda_\sigma)_\sigma, \text{ where } \lambda_\sigma = -n_\sigma (m_\sigma + 2), \text{ for } m_\sigma \in \mathbb{N}_{\geq 0}.
\end{equation}
These requirements imply that $\varphi$ is identified with a smooth, $K_{\mathcal{O}'} K_\infty$-invariant function, contained in the subspace $V_\varphi \subset L^2(G(F) \backslash G(\mathbb{A}))$ of an automorphic representation $\pi = \otimes_v \pi_v$ (as before we may assume that $\pi$ is infinite dimensional).

By [HS12, Thm. 4.13], there exists an automorphic representation $\tilde{\pi} \simeq \otimes_v \tilde{\pi}_v$ of $G(\mathbb{A})$ (with unitary central character) such that $V_{\pi} \subset V_{\tilde{\pi}}^{1,1} G(\mathbb{A})$ where $V_{\pi}^{1,1} G(\mathbb{A})$ denotes the restriction to $G(\mathbb{A})$ of a certain subspace of $V_\pi$ (in most cases $V_{\pi}^{1,1} = V_\pi$). In other words, $\tilde{\pi}$ is an automorphic representation on $B^\times(\mathbb{A}_E)$ whose central character on $Z_E(\mathbb{A}) \simeq \mathbb{A}_E^\times$ is trivial when restricted to $Z_F(\mathbb{A}) \simeq \mathbb{A}_E^\times$. 
In particular, in order to bound \( \varphi \) it is sufficient to bound \( \tilde{\varphi} \), since obviously

\[
\|\varphi\|_\infty \lesssim \|\tilde{\varphi}\|_\infty,
\]

but we also need to compare the \( L^2 \)-norms. We proceed to show that \( \|\varphi\|_2 \) and \( \|\tilde{\varphi}\|_2 \) are of comparable size. First observe that since \( \tilde{\varphi} \) is invariant by an open compact subgroup of \( \widehat{G}(A_f) \) and by the subgroup \( \{ w \in \widehat{\mathcal{O}}_{B_E} \mid nr(w) \in \mathcal{X}_\infty \} \), it is invariant under \( \widehat{\mathcal{O}}_{B_E}^\times \). In particular \( \pi_v \) is unramified for every finite place not dividing \( nr_F(disc \mathcal{O}_{B_E}) \). It follows from this and [HS12, Rem. 4.20] that for any choice for Haar measures

\[
\frac{\langle \varphi, \varphi \rangle_G}{\langle \tilde{\varphi}, \tilde{\varphi} \rangle_{\widehat{G}}} = C(\tilde{\pi}) \frac{\text{vol}(G(F) \setminus G(\mathbb{A}))}{\text{vol}(Z_E(\mathbb{A}) G(F) \setminus G(\mathbb{A}))}
\]

where \( C(\tilde{\pi}) = nr_F(disc \mathcal{O}_{B_E})^{o(1)} \). We fix the Haar measure compatibly with the fact that \( G \) is the kernel of the \( F \)-algebraic map

\[
\widehat{G} \rightarrow \mathbb{G}_{m,E}^1 = \{ x \in \mathbb{G}_{m,E} \mid nr_{F/E} x = 1 \}
\]

by fixing a Haar measure on \( \mathbb{G}_{m,E}^1(A) \), and the Haar measure on \( PB^\times(\mathbb{A}_E) \) so that \( PB^\times(\mathbb{E}_E) \) and the image of \( \widehat{\mathcal{O}}_{B_E}^\times \) in \( PB^\times(\mathbb{A}_F) \) have volume 1. It follows from these choices that

\[
\text{vol}(Z_E(\mathbb{A}) G(F) \setminus G(\mathbb{A})) = nr_E(disc \mathcal{O}_{B_E})^{1/2 + o_E(1)},
\]

\[
\text{vol}(G(F) \setminus G(\mathbb{A})) = vol(X(3)(\mathcal{O})) = nr_E(disc \mathcal{O})^{1 + o_E(1)} = nr_E(disc \mathcal{O}_{B_E})^{1/2 + o_E(1)},
\]

and hence

(6.3) \[
\frac{\langle \varphi, \varphi \rangle_G}{\langle \tilde{\varphi}, \tilde{\varphi} \rangle_{\widehat{G}}} = nr_E(disc \mathcal{O}_{B_E})^{o(E(1)}.
\]

6.2. The archimedean places. The algebra \( E \) splits above the infinite places of \( F \): \( E_\infty = \mathbb{R} \times \mathbb{R} \) and \( B_{E_\infty}^\times = B_{\mathbb{R}}^\times \times B_{\mathbb{R}}^\times \). By hypothesis \( \tilde{\pi}_\sigma \) admits for each \( \sigma \) a non-zero \( K_\sigma \)-invariant vector. This implies that as a representation of \( B^\times(\mathcal{O}_E) \) we have

\[
\tilde{\pi}_\sigma \simeq \Pi_{m_\sigma} = \pi_{m_\sigma} \boxtimes \pi_{m_\sigma},
\]

where \( m_\sigma \) is defined by (6.2). For such a representation, the space of \( K_\sigma \simeq \text{SO}_3(\mathbb{R}) \) invariant vectors is one-dimensional. Let \( E_{m_\sigma} \) be such a non-zero vector, and for \( g = (g_1, g_2) \in B_\sigma^1 \times B_\sigma^1 \) let

\[
P_{m_\sigma}(g_\sigma) = \frac{\langle g_\sigma.E_{m_\sigma}, E_{m_\sigma} \rangle}{(E_{m_\sigma}, E_{m_\sigma})}
\]

be the corresponding matrix coefficient. For \( g = (g_\sigma)_{\sigma} \in B^1(E_\infty) = B^1(F_\infty) \times B^1(F_\infty) \) let

\[
P_m(g) = \prod_{\sigma} P_{m_\sigma}(g_\sigma)
\]

be the product of these matrix coefficients.

6.3. The pre-trace formula. Let us (for notational simplicity) first assume that \( E \) is a field. By the amplified pretrace formula for \( B^\times(\mathcal{O}_E) \) we have for \( g \in B^\times(\mathcal{O}_E) \) and a suitable parameter \( L \) to be chosen in a moment

(6.4) \[
\frac{|\tilde{\varphi}(g)|^2}{\langle \tilde{\varphi}, \tilde{\varphi} \rangle_{\widehat{G}}} \ll |m|^2 L^\varepsilon \left( \frac{1}{L} + \sum_{i=1}^4 \frac{1}{L^{2+i/2}} \sum_{\ell \in \mathcal{L}} \sum_{\gamma \in \mathcal{O}_{B_E^\times}/nr(\gamma)=\ell} |P_m(g^{-1} g_\gamma)| \right).
\]
This is the same expression as (5.1) except that the underlying field is now called $E$ rather than $F$ (this applies also to the definition of the sets $\mathcal{L}_i$), and the matrix coefficient is different. In particular, $\mathcal{O}_{B_E, g}$ is locally everywhere conjugated to $\mathcal{O}_{B_E}$, and its level $\pi_{B_E}$ satisfies

$$\pi_{E}(\pi_{B_E}) \asymp E \pi_{E}(\text{disc}\mathcal{O}_{B_E})^{1/2}$$

as in Section 5.2. Let $V = \pi_{E}(\text{disc}\mathcal{O}_{B_E})^{1/2} = \pi_{E}(\mathcal{O}_E)^{1+o(1)}$. Now the trivial bound $|P_m(g)| \leq 1$ together with the argument of Section 5.2 shows

$$|\tilde{\varphi}(g)| \ll \varepsilon \min |V^{-1/6+\varepsilon}\|\varphi\|$.

In view of (6.2) and (6.3) this may be rewritten

$$\|\varphi\|_\infty \ll \varepsilon |\chi|^{1/2} \text{vol}((X_3(\mathcal{O}))^{-1/6+\varepsilon}$$

if $\|\varphi\|_2 = 1$, which matches the generic bound in the $\lambda$-aspect and improves it in the volume aspect.

The split case $E = F \times F$ is very similar. In this case the sums over $\ell \in \mathcal{L}_i$ and over $\gamma$ in (6.4) factor, and we can apply the argument of Section 5.2 for each factor. This completes the proof of Theorem 2.

**Remark 6.1.** The function $P_m$ is bi-$\Delta B^1_\sigma$ invariant (i.e. spherical) and may be expressed in terms of the character $\chi_m$ of the representation $\pi_m$, namely

$$P_m(g) = P_m(g_1/g_2, 1) = \chi_m(t) = \frac{1}{m+1} \sin((m+1)\theta) \sin(\theta),$$

$$t = \cos(\theta) = \langle g_1/g_2, 1 \rangle_B = \frac{1}{2} \text{tr}(g_1/g_2).$$

In particular, one has

$$P_m(g_1, g_2) \leq \min \left(1, \frac{1}{(m+1)(1-t^2)^{1/2}} \right).$$

Just as for Lemma 1, this decay property (as $t$ gets away from $\pm 1$) may be used together with an appropriate version of Lemma 5 to yield a hybrid bound for $\varphi$. We leave this to future work.

**References**


