SUBCONVEXITY FOR SUP-NORMS OF CUSP FORMS ON PGL(n)

VALENTIN BLOMER AND PÉTER MAGA

ABSTRACT. Let F be an L^2-normalized Hecke Maaß cusp form for $\Gamma_0(N) \subseteq \mathrm{SL}_n(\mathbb{Z})$ with Laplace eigenvalue λ_F. If Ω is a compact subset of $\Gamma_0(N)\backslash \mathrm{PGL}_n/PO(n)$, we show the bound $\|F|_\Omega\|_\infty \ll_\Omega N^\varepsilon \lambda_F^{n(n-1)/8-\delta}$ for some constant $\delta = \delta_n > 0$ depending only on n.

1. INTRODUCTION

1.1. **The main result.** Given a Riemannian locally symmetric space $X = \Gamma\backslash S$, it is a classical question in analysis to find pointwise bounds for eigenfunctions $F \in L^2(X)$ of the algebra $\mathscr{D}(S)$ of invariant differential operators, uniformly in X, in terms of their Laplacian eigenvalue λ_F. If X is a compact locally symmetric space of rank r, then Sarnak [Sa1] proved that an L^2-normalized joint eigenfunction F satisfies

$$\|F\|_\infty \ll \lambda_F^{(\dim X - r)/4}. \tag{1.1}$$

This is often referred to as the *convexity bound*, and it is sharp in general. The proof comes only from local considerations and uses, among other things, various properties and asymptotics of spherical functions. The same proof works for non-compact spaces X, provided F is restricted to a compact domain, but it was observed recently by Brumley and Templier [BT] that (1.1) is wrong in general, for instance in the case $X = \mathrm{PGL}_n(\mathbb{Z})\backslash \mathrm{PGL}_n(\mathbb{R})/\mathrm{PO}_n$ if $n \geq 6$, and probably for many other spaces of high rank, too.

There are many more refined versions and conjectures on the sup-norm problem. Sarnak's *purity conjecture* [Sa1] states that the accumulation points of $\log \|F\|_\infty / \log \lambda_F$ are contained in $\frac{1}{4}\mathbb{Z} \cap [0, (\dim X - r)/4]$. We mention here in particular the case of hyperbolic 3-space with $\dim X = 3$, $r = 1$, where theta lifts produce eigenfunctions with $\|F\|_\infty \gg \lambda_F^{1/4}$ [RS]. For spaces of negative curvature one expects that (1.1) is not sharp and stronger bounds hold true, although this has not yet been proved in any particular case. In this paper we are interested in *arithmetic* situations: many classical examples of Riemannian locally symmetric spaces enjoy additional symmetries given by the Hecke operators, a commutative family of normal operators, and the arithmetically interesting functions on this space are not only eigenfunctions of $\mathscr{D}(S)$, but in addition joint eigenfunctions of the Hecke algebra $\mathscr{H}$. The *subconvexity conjecture* predicts an upper bound with an exponent strictly smaller than $(\dim X - r)/4$ for joint eigenfunctions of $\mathscr{D}(S)$ and the Hecke algebra $\mathscr{H}$, at least on compact spaces or when restricted to compact domains of non-compact spaces.

An important motivation for the large eigenvalue limit comes from the correspondence principle of quantum mechanics. In the situation of compact Riemannian manifolds of negative curvature, the quantum unique ergodicity conjecture [RS] asserts that all eigenfunctions become equidistributed in terms of measure convergence. A different, but not unrelated measure of equidistribution is given by $\|F\|_\infty$, a quantitative version of which is a subconvex bound over (1.1). In addition, subconvex bounds for $\|F\|_\infty$ for joint eigenfunctions F have diverse analytic and – in arithmetic situations – number theoretical applications, of which we only mention the multiplicity problem [Sa1], control over the zero set or nodal lines of automorphic

2010 *Mathematics Subject Classification.* 11F55, 11F72, 11D75.

Key words and phrases. sup-norms, Hecke operators, trace formula, diophantine approximation, amplification, GL(n).

The first author was supported by the Volkswagen Foundation and Starting Grant 258713 of the European Research Council. The second author was supported by Starting Grant 258713 of the European Research Council and OTKA grant no. NK104183.

forms [Ru, GRS], and number theoretic investigations of Hecke eigenvalues, in particular in connection with L-functions and shifted convolution problems [BH, HM, Mag].

The first breakthrough in the subconvexity problem for sup-norms of automorphic forms was achieved by Iwaniec and Sarnak [IS] in the classical situation $X = \Gamma\backslash\mathcal{H}_2$ where $\mathcal{H}_2$ is the hyperbolic plane and $\Gamma \leq \mathrm{SL}_2(\mathbb{R})$ is a cocompact arithmetic subgroup or $\mathrm{SL}_2(\mathbb{Z})$. For L^2-normalized Hecke Maaß cusp forms F they proved the bound $\|F\|_\infty \ll \lambda_F^{5/24+\varepsilon}$. Other rank one cases include congruence quotients of hyperbolic 3-space ([BHM]). Up until recently, however, no higher rank examples were known, and only very recently the subconvexity conjecture for sup-norms has been solved for automorphic forms for the groups $\mathrm{Sp}_4(\mathbb{Z})$ [BP], $\mathrm{SL}_3(\mathbb{Z})$ [HRR] and $\mathrm{SL}_4(\mathbb{Z})$ [BM].

As discussed in [BM], the subconvexity problem for sup-norms has not only the name in common with the subconvexity problem for L-functions, but it also shares methodological features and in particular the fact that there is a considerable history of results for subgroups of $\mathrm{SL}_2(\mathbb{Z})$, but only very few sporadic results have recently become available in situations of small rank > 1. Unfortunately standard and even the most advanced techniques from analytic number theory often fail to be powerful enough in situations of unbounded rank.

In this article we solve for the first time the subconvexity problem for sup-norms of automorphic forms on

$$\Gamma\backslash G/K, \quad G = \mathrm{PGL}_n(\mathbb{R}), \quad K = \mathrm{PO}_n, \quad \Gamma = \Gamma_0(N)$$

for *arbitrary* n where $\Gamma_0(N)$ is the usual congruence subgroup of $\mathrm{SL}_n(\mathbb{Z})$ with bottom row congruent to $(0, \dots, 0, *)$ modulo N. The symmetric space G/K has dimension $(n-1)(n+2)/2$ and rank $n-1$, hence the convexity exponent is $n(n-1)/8$. We equip $\Gamma\backslash G/K$ with an inner product in a way that $\mathrm{vol}(\Gamma\backslash G/K) = [\mathrm{SL}_n(\mathbb{Z}) : \Gamma] = N^{n-1+o(1)}$. We will explain the new ingredients in detail in the next subsection and proceed with the statement of our main result. Let $W \cong S_n$ be the Weyl group, A be the diagonal torus in G and $\mathfrak{a}$ the corresponding Lie algebra. Let Σ denote the set of roots of $\mathfrak{g}$, and fix a maximal subset Σ^+ of positive roots. For $\alpha \in \Sigma$, let $m(\alpha)$ denote the dimension of the corresponding root space. The Killing form defines a (W-invariant) inner product on $\mathfrak{a}$ which induces an inner product on $\mathfrak{a}^*$ via root vectors. For $\lambda \in \mathfrak{a}^*$ define

$$D(\lambda) = \prod_{\alpha \in \Sigma^+} (1 + |\langle \alpha, \lambda \rangle|)^{m(\alpha)}. \tag{1.2}$$

An eigenfunction F of $\mathscr{D}(G/K)$ has a spectral parameter $\mu \in \mathfrak{a}_{\mathbb{C}}^*$ such that $\mathcal{D}F = \gamma(\mathcal{D})(\mu)F$ for all $\mathcal{D} \in \mathscr{D}(G/K)$, where γ is the Harish-Chandra homomorphism.

Theorem 1. *Let $n \geq 2$. Let F be a Hecke Maaß cusp form for $\Gamma_0(N)$ with spectral parameter $\mu \in \mathfrak{a}_{\mathbb{C}}^*$, and write $\mu^* := \Re\mu \in \mathfrak{a}^*$. Let Ω be a fixed compact subset of G/K. Then*

$$\|F|_\Omega\|_\infty \ll_{\Omega,\varepsilon} N^\varepsilon D(\mu^*)^{1/2-\delta} \tag{1.3}$$

for some (effectively computable) constant $\delta = \delta_n > 0$ and any $\varepsilon > 0$. The implied constant depends at most on Ω and ε. In particular, if λ_F denotes the Laplacian eigenvalue of F, then

$$\|F|_\Omega\|_\infty \ll_{\Omega,\varepsilon} N^\varepsilon \lambda_F^{\frac{n(n-1)}{4}(\frac{1}{2}-\delta)}. \tag{1.4}$$

We emphasize that this result holds for *all* Hecke Maaß forms without any assumptions on the Ramanujan conjecture, neither at finite places nor at the infinite place. Although strictly positive, our proof produces only a very small value of δ_n that does not yield any information on the next eligible exponent in Sarnak's purity conjecture. As mentioned before, the restriction to a compact subset Ω is a necessary condition in view of [BT].

Under the Ramanujan conjecture at infinity, Theorem 1 was proved independently by Simon Marshall in the important preprint [Mar] in the more general context of split semisimple groups by quite different diophantine techniques for the estimation the number of Hecke returns (partly based on unpublished notes of Sarnak-Venkatesh). We believe that both techniques are of independent interest, and it would be interesting to compare them more thoroughly.

1.2. **Counting techniques.** The presence of Hecke operators transforms the purely analytic problem of bounding eigenfunctions on manifolds into a problem that has an intersection with several branches of mathematics, in particular number theory. The starting point is the spectral expansion of an automorphic kernel: we consider a weighted spectral sum

$$\sum_{\varpi} A(\varpi)|F_\varpi(g)|^2$$

over the constituents ϖ of $L^2(\Gamma\backslash G/K)$ (including Eisenstein series, so that the sum is in reality a combination of sums and integrals) where $A(\varpi)$ is a non-negative weight function with $A(\varpi_0) = 1$ for the specific cuspidal automorphic representation ϖ_0 whose sup-norm we want to bound (and $A(\varpi)$ small otherwise). Dropping all but one term, we recover a bound for $F_{\varpi_0}(z)$. A general amplifier $A(\varpi)$ for GL(n) has been constructed in [BM, Section 4] and consists of double cosets

$$\Gamma(p^\nu, 1, \dots, 1)\Gamma \quad \text{and their adjoints} \quad \Gamma(p^\nu, \dots, p^\nu, 1)\Gamma$$

for $1 \le \nu \le n$. The geometric side of the spectral expansion of the automorphic kernel features a diophantine problem which in all treatments of the subconvexity problem for sup-norms is the heart of the matter and reflects the arithmeticity of the underlying problem. In the case of GL(n), one has to count matrices $\gamma \in \mathrm{Mat}(n, \mathbb{Z})$ satisfying

$$\gamma^\top Q\gamma = (\det\gamma)^{2/n}Q + \text{very small error} \tag{1.5}$$

where $Q \in \mathrm{Mat}(n, \mathbb{R})$ is a fixed positive definite matrix depending on the point $g \in G/K$ at which we want to bound F_{ϖ_0}. The choice of our amplifier (see (2.5)) leads to matrices γ with

$$\det\gamma = q^\nu p^{\nu(n-1)} \tag{1.6}$$

for $1 \le \nu \le n$ and primes $q, p \asymp L$ of the same order of magnitude. Moreover, if $\Delta_j(\gamma)$ denotes the j-th determinantal divisor, i.e. the greatest common divisor of all j-by-j minors, then

$$\Delta_2(\gamma) = p^\nu. \tag{1.7}$$

This condition means roughly that any two columns of γ are multiples of each other modulo p^ν. It turns out that we have to show that the number of $\gamma \in \mathrm{Mat}(n, \mathbb{Z})$ satisfying (1.5) – (1.7) is

$$O(L^{\nu(n-1)-\delta}) \tag{1.8}$$

for some $\delta > 0$. Solving the counting problem (1.5) – (1.8) in full generality is the most novel part of this paper for which several new ideas are necessary that we proceed to describe here in informal language for the reader's convenience.

If $Q = \mathrm{id}$ is the identity matrix, the argument is fairly simple: let $\gamma_1, \dots \gamma_n \in \mathbb{Z}^n$ denote the columns of γ. We distinguish three cases.

Case 1: If $q \ne p$ and $2\nu/n \notin \mathbb{N}$, then the left hand side of (1.5) is integral, but the right hand side is not, at least if the error is sufficiently small. Therefore there are no solutions at all in this case.

Case 2: If $q \ne p$, but $2\nu/n \in \mathbb{N}$, then we write $\gamma_1 \equiv a\gamma_2 \pmod{p^\nu}$, and substituting this into $\langle\gamma_1, \gamma_2\rangle = 0$, $\|\gamma_1\|^2 = \|\gamma_2\|^2 = (qp^{n-1})^{2\nu/n}$, one obtains the congruence $1 + a^2 \equiv 0 \pmod p$. If we restrict to primes $\equiv 3 \pmod 4$, this leads to a contradiction, too.

Case 3: If $q = p$, we choose the first column γ_1 of γ randomly. Its n entries satisfy a quadratic equation by (1.5), so there are at most $O(L^{\nu(n-2+\varepsilon)})$ choices for γ_1. Comparing with (1.8), almost everything else should now be determined. It is not hard to see that (1.5) – (1.7) imply that in this case any two choices γ_2, γ_2' for the second column satisfy $\langle\gamma_2, \gamma_2'\rangle \equiv 0 \pmod{p^{2\nu}}$, and since $\|\gamma_2\| = \|\gamma_2'\| = p^\nu$, this means that γ_2 and γ_2' are either parallel or orthogonal, hence there are $O(1)$ choices for γ_2 and analogously for all other columns $\gamma_3, \dots, \gamma_n$. (Perhaps higher Δ_j's could also be implemented into the above diophantine analysis to improve the value of δ in (1.8).)

A similar argument works if Q is (very close to) a rational matrix of small height L^ε, and we now describe a (doubly) recursive strategy to achieve a situation where Q is a rational matrix of small height. It is based on two ideas that to our knowledge have not yet been applied in the amplification method: (a) we have the flexibility to vary L – maybe some ranges are better suited than others, and (b) we show that there exists a matrix Q' with rational or at least algebraic entries of not small, but controllable height, with the property that every γ satisfying (1.5) – (1.7) also satisfies (1.5) with Q' in place of Q. In other words, for the purpose of counting solutions to (1.5) – (1.7), we can exchange Q for Q', and the latter has better diophantine properties. To be more precise, consider the operator

$$B_\gamma : Q \mapsto \gamma^\top Q\gamma - (\det\gamma)^{2/n}Q.$$

For an admissible γ, the matrix Q is close to $\ker(B_\gamma)$, hence Q is close to the subspace

$$H_0 := \bigcap_{\substack{\gamma \text{ satisfying (1.5) – (1.7)}\\ q,p\in I_0=[L,2L]}} \ker(B_\gamma).$$

By definition, any matrix $Q' \in H_0$ has the property that all admissible γ for Q are also admissible for Q', in fact with *no error term* in (1.5). This type of "rigidity" is often a key ingredient in such counting problems. Now we repeat this procedure but for the larger intervals $q, p \in I_j := [L, 2L^{D^j}]$, $j = 1, 2, \ldots$, getting a chain of finite-dimensional vector spaces $H_0 \supseteq H_1 \supseteq H_2 \supseteq \ldots$. At some point we must have $H_i = H_{i+1}$. Then any γ solving (1.5) for primes $q, p \in I_{i+1}$ also solves (1.5) when Q is replaced with an arbitrary $Q' \in H_i = H_{i+1}$ without error term. Now H_i is defined over an algebraic number field containing n-th roots of primes in I_j, and restricting our attention to primes in $I_{i+1} \setminus I_i$, we arrive at a contradiction in case 1 above. Now we run a second version of this recursive argument inside the interval $I_{i+1} \setminus I_i$ and restrict ourselves to the cases 2 and 3 above where $(qp^{n-1})^{2\nu/n}$ is an integer. We choose again a chain of strongly increasing intervals $I_0' \subseteq I_1' \subseteq \ldots$ and obtain a corresponding chain of spaces $H_0' \supseteq H_1' \supseteq \ldots$ that in this case are defined over $\mathbb{Q}$. When $H_k' = H_{k+1}'$, we choose a rational matrix $Q' \in H_k'$ of controlled height. However, with respect to the primes in the larger interval I_{k+1}', this height is very small, and we can proceed as described in the previous paragraph.

This technique, carried out in detail in Sections 3 – 6 works in much greater generality. In particular, it is not restricted to the group $\mathrm{GL}(n)$ and can be applied in different amplification settings. In the present situation it provides the additional insight that the solution of the arithmetic sup-norm problem is determined by the points that behave roughly like the identity in terms of diophantine approximation.

Finally we mention an important technical point: Case 2 requires us to consider subsets of primes that satisfy certain quadratic residue properties. Although the primes in I_{k+1}' typically are much larger than the primes in I_k', they are only polynomially larger, and this is outside the range of Siegel-Walfisz type theorems. Instead we need quantitative versions of Linnik type results on primes in arithmetic progressions. Hence our argument uses implicitly log-free density theorems for Dirichlet L-functions and the Deuring-Heilbronn phenomenon.

We hope that these remarks will guide the reader through the proof of Theorem 1. We end the introduction by thanking the referee for a very careful reading of the manuscript that improved the exposition.

2. An amplified pretrace formula

The description of the set-up is similar as in [BM], but we take care to avoid the Ramanujan conjecture. Let $C : G \to \mathfrak{a}/W$ be the Cartan projection, so that

$$g = k_1 \exp(C(g))k_2 \tag{2.1}$$

with $k_1, k_2 \in K$. The half-sum $\rho \in \mathfrak{a}^*$ of positive roots is given by

$$\rho = \frac{1}{2}\sum_{j=1}^{n}(n+1-2j)e_j \in \mathfrak{a}^*$$

where $e_j(\mathrm{diag}(a_1,\dots,a_n)) = a_j$. As usual, we denote by C_ρ the convex hull of the points $\{w\rho \mid w \in W\}$.

Let $F \in L^2(\Gamma\backslash G/K)$ be a Hecke Maaß cusp form which we view both as a function on G/K and a right K-invariant function on G. At the archimedean place, it comes with a spectral parameters $\mu = (\mu_1,\dots,\mu_n) \in \mathfrak{a}_{\mathbb{C}}^*/W$ in the set

$$\Lambda := \Big\{\mu \in \mathfrak{a}_{\mathbb{C}}^*/W \mid \sum_{j=1}^{n}\mu_j = 0,\ \{\mu_1,\dots,\mu_n\} = \{\bar\mu_1,\dots,\bar\mu_n\},\ \mu \in \mathfrak{a}^* + iC_\rho\Big\}. \tag{2.2}$$

(Better bounds for the imaginary parts of μ_j are available, but we do not need stronger results.) As in Theorem 1 we write

$$\mu^* := \Re\mu.$$

For the proof of Theorem 1 we may and will assume from now on that $\|\mu\|$ and hence $\|\mu^*\|$ are sufficiently large. The Laplacian of F is given by

$$\lambda_F = \frac{n^3-n}{24} + \frac{1}{2}(\mu_1^2 + \ldots + \mu_n^2) \asymp \|\mu^*\|^2.$$

The Harish-Chandra $\mathbf{c}$-function satisfies

$$\frac{1}{|\mathbf{c}(\lambda)|^2} \asymp \prod_{1\le j<k\le n}|\lambda_j-\lambda_k|\tanh(|\lambda_j-\lambda_k|) \ll D(\lambda) \ll 1+\|\lambda\|^{n(n-1)/2} \tag{2.3}$$

for $\lambda \in \mathfrak{a}^*/W$ and $D(\lambda)$ as in (1.2). In particular, (1.4) is an immediate consequence of (1.3).

Next we set up the spectral expansion of an automorphic kernel. For our choice of test function we need the following lemma which allows us to avoid the Ramanujan conjecture at infinity.

Lemma 1. *Let $c > 0$. There exists an even function $h : \mathbb{C} \to \mathbb{C}$ with compactly supported Fourier transform such that $h(\mathbb{R}) \subseteq [0,\infty)$, $\Re h|_{\{|\Im z|\le c\}} \ge 0$ and $\Re h|_{\{|z|\le c\}} \ge 1$.*

Proof. We start with the function

$$g(x) = \frac{\sin(\pi x)}{\sin(\pi\sqrt{x})\sinh(\pi\sqrt{x})}.$$

It is obviously holomorphic outside the non-positive real axis (the possible poles at integral squares cancel). In addition, it is even and has Taylor series at 0, hence g is entire. Moreover, $g(x) \ll e^{\pi|\Im x|}$, so that by the Paley-Wiener theorem the Fourier transform of g is compactly supported. Next we define

$$G(y) = y\cdot g(y)^2, \quad h(x) = -\int_{-\infty}^{x} G(y)\,dy.$$

Then h is a non-negative, even function again with compactly supported Fourier transform that is monotonically increasing for $x < 0$ and decreasing for $x > 0$. One checks that

$$G(m^2) = \frac{4m^4}{\sinh(\pi m)^2} \quad (m \in \mathbb{N}), \qquad G(x+it) \ll \frac{x^2}{\sinh(\pi\sqrt{x})^2} \quad (x \ge 1, |t| \le 1),$$

so that by Cauchy's integral formula also

$$G'(x+it) \ll \frac{x^2}{\sinh(\pi\sqrt{x})^2} \quad (x \ge 2, |t| \le 1/2).$$

In particular, there exists a constant $C>0$ such that $G(x)\gg m^4\sinh(\pi m)^{-2}$ for $|x-m^2|\le C$ so that

$$|h'(x+it)|=|G(x+it)|\ll\frac{x^2}{\sinh(\pi\sqrt{x})^2}\ll\int_{\lceil\sqrt{x}\rceil^2}^{\lceil\sqrt{x}\rceil^2+1}G(x)\le h(x)$$

for $x\ge 2$, $|t|\le 1/2$. We conclude that $\Re h$ is non-negative in some fixed horizontal strip and bounded below in some fixed ball about 0. Rescaling $h(x)$ as $\alpha h(x/\beta)$ for suitable $\alpha,\beta>0$, we conclude the claim.

Identifying $\mathfrak{a}^*$ with the trace zero hyperplane in $\mathbb{R}^n$, we now define

$$f:\mathfrak{a}_{\mathbb{C}}^*\to\mathbb{C},\quad(\lambda_1,\dots,\lambda_n)\mapsto h(\lambda_1)\cdot\ldots\cdot h(\lambda_n)$$

where h is as in the previous lemma, so that f is a fixed function on $\mathfrak{a}_{\mathbb{C}}^*$ with compactly supported Fourier transform such that f is real on $\mathfrak{a}^*$, $\Re f$ is non-negative in the strip $|\Im\lambda_j|\le\|\rho\|_\infty$, say, and $\Re f\ge 1$ on a ball in $\mathfrak{a}_{\mathbb{C}}^*$ about 0 of radius $\|\rho\|_2$. We define

$$\tilde{f}_\mu(\lambda):=\Big(\sum_{w\in W}f(\mu^*-w\cdot\lambda)\Big)^2.$$

This has again compactly supported Fourier transform, and the support is independent of μ. One verifies quickly that

$$\tilde{f}_\mu(\lambda)\ge 0$$

for all $\lambda\in\Lambda$ as defined in (2.2) and

$$\tilde{f}_\mu(\mu)\ge 1.$$

Both estimates are consequences of the unitarity condition in (2.2) and the fact that $f(\lambda)+f(\bar{\lambda})=2\Re f(\lambda)$, so that we can use the lower bounds for $\Re f$. Moreover, the rapid decay along the real axis shows

$$\tilde{f}_\mu(\lambda)\ll_A\max_{w\in W}(1+\|\mu^*-w\cdot\lambda\|)^{-A}$$

for $\lambda\in\mathfrak{a}^*$ and any $A>0$. By the Harish-Chandra inversion formula together with the uniform bounds for elementary spherical functions in [BP, Theorem 2], we see that the inverse spherical transform $f_\mu: K\backslash G/K\to\mathbb{C}$ of $\tilde{f}_\mu$ has compact support and satisfies the decay property

$$f_\mu(g)\ll D(\mu^*)\big(1+\|\mu^*\|\|C(g)\|\big)^{-1/2},\tag{2.4}$$

cf. [BM, (3.9)].[1]

Now let $L_0>5$ and let $\mathcal{P}$ be a set of primes in $[L_0,2L_0]$ coprime to N. For $m,l\in\mathbb{N}$ define

$$S(m,l):=\{\gamma\in\mathrm{Mat}(n,\mathbb{Z})\mid\det\gamma=m,\ \Delta_1(\gamma)=1,\ \Delta_2(\gamma)=l\}$$

where as in the introduction $\Delta_j(\gamma)$ denotes the j-th determinantal divisor. With this notation it has been shown in [BM, (6.2)] that[2]

$$|\mathcal{P}|^2|F(g)|^2\ll|\mathcal{P}|D(\mu^*)+\sum_{\nu=1}^n\sum_{p,q\in\mathcal{P}}\frac{1}{L_0^{(n-1)\nu}}\sum_{\gamma\in S(q^\nu p^{(n-1)\nu},p^\nu)}|f_\mu(g^{-1}\gamma g)|\tag{2.5}$$

for $g\in G$. This has been shown for cuspidal automorphic forms F for $\mathrm{SL}_n(\mathbb{Z})$, but it holds verbatim for the congruence subgroup $\Gamma_0(N)$, as long as we avoid ramified Hecke operators. In fact, the counting problem becomes even easier as the matrices γ counted in $S(m,l)$ have to satisfy additional congruence properties. For the purpose of getting upper bounds, we can ignore these extra conditions.

Fix some large $M>1$. Using the notation (2.1), we write $C_{\gamma,g}:=\|C(g^{-1}\gamma g)\|$. Since f_μ has compact support, only those γ with $C_{\gamma,g}\ll 1$ contribute to the sum (2.5). The contribution of γ with $C_{\gamma,g}\ge L_0^{-M}$ is small because of the decay property (2.4) of the function f_μ. For the remaining γ we estimate the function

[1] We remark that in [BM, (3.9)] and similar displays $\mathbf{c}(\mu)$ should be replaced with $D(\mu)^{-1/2}$.

[2] In [BM] the set $\mathcal{P}$ was the set of all primes, but the argument works verbatim for any set of primes.

f_μ trivially by $D(\mu^*)$ and need good bounds for the number of such matrices occurring in the sum (2.5). They satisfy $\gamma^\top Q\gamma = (\det\gamma)^{2/n}Q + O((\det\gamma)^{2/n}L_0^{-M})$ where

$$Q = (\det g)^2 g^{-\top}g^{-1} = (Q_{ij}) \in \mathrm{Mat}_n(\mathbb{R}) \tag{2.6}$$

is a positive definite symmetric matrix. With this in mind, define

$$\mathcal{S}(Q,a,b,M) := \left\{\gamma \in \mathrm{Mat}_n(\mathbb{Z}) \mid \gamma^\top Q\gamma = (ab^{n-1})^{2/n}Q + O((ab^{n-1})^{(2-M)/n}),\ \Delta_1(\gamma)=1,\ \Delta_2(\gamma)=b\right\} \tag{2.7}$$

for $a,b \in \mathbb{N}$ and $M>0$. We also formally allow $M=\infty$ in which case there is no error term. Following the argument in [BM, Section 6, see in particular display after (6.5)], we obtain the basic estimate

$$|F(g)|^2 \ll D(\mu^*)\left(\frac{1}{|\mathcal{P}|} + D(\mu^*)^{-\frac{1}{n(n-1)}}L_0^{n^3+M/2} + \sum_{\nu=1}^{n}\frac{1}{|\mathcal{P}|^2}\sum_{p,q\in\mathcal{P}}\frac{\mathcal{S}(Q,q^\nu,p^\nu,M)}{L_0^{\nu(n-1)}}\right). \tag{2.8}$$

It is now clear that we have to bound the cardinality of $\mathcal{S}(Q,q^\nu,p^\nu,M)$ which is the counting problem discussed in the introduction. The next four sections are devoted to this task. We remark that this argument uses crucially that $g \subseteq G/K$ and hence Q are in some fixed compact domain, so that for instance the implicit constant in (2.7) is independent of Q.

3. Auxilliary lemmas

Lemma 2. *Let Q be a symmetric, positive definite matrix, $\varepsilon>0$. Then*

$$|\{y\in\mathbb{Z}^n \mid y^\top Qy = m^2\}| \ll_Q m^{n-2+\varepsilon}$$

for all $m\geq 1$ (where the implied constant can be chosen as a continuous function of the successive minima of Q).

Proof. This follows from the special case $k=0$, $\delta=0$, $q_0=m^2$ of [BM, Corollary 5.3], but for convenience we repeat the argument. We can assume that Q is Minkowski-reduced ([Ca, Chapter 12]) with successive minima $h_1\leq\ldots\leq h_n$. For $j=n,n-1,\ldots,3$ we can choose successively y_j in $O(1+mh_j^{-1/2})$ ways. We are then left with an inhomogeneous binary problem in y_1,y_2 whose (positive definite) quadratic homogeneous part has discriminant $|D|\gg h_1^2$. By [BP, Corollary 9] with $\delta=0$ there are at most $\ll_{h_1,h_n} m^\varepsilon$ choices for y_1,y_2.

We will use the following lemma to exploit the determinantal condition (1.7) for two columns x, y of γ. We call an integral vector $x\in\mathbb{Z}^n$ completely divisible by an integer m if all its entries are divisible by m. The letters p and q are reserved for prime numbers.

Lemma 3. *Let p be a prime, $\rho\in\mathbb{N}$. Let $x=(\xi_1,\ldots,\xi_n)$, $y=(\eta_1,\ldots,\eta_n)\in\mathbb{Z}^n$ be two integral vectors satisfying*

$$\xi_i\eta_j \equiv \xi_j\eta_i \pmod{p^\rho}$$

for $1\leq i,j\leq n$. Assume that both vectors are not completely divisible by p. Let $A=A^\top\in\mathrm{Mat}_n(\mathbb{Z})$. Then the following holds.

(a) There exists (a unique) $a\in(\mathbb{Z}/p^\rho\mathbb{Z})^*$ such that $y\equiv ax \pmod{p^\rho}$.

(b) With a as in part (a), we have

$$2x^\top Ay \equiv a\cdot x^\top Ax + \bar{a}\cdot y^\top Ay \pmod{p^{2\rho}}. \tag{3.1}$$

Proof. Assume without loss of generality that $p\nmid\xi_n$. We show $p\nmid\eta_n$. Indeed, assume the contrary. Then by assumption there exists an index $j\neq n$ such that $p\nmid\eta_j$, but this is a contradiction to $0\not\equiv\xi_n\eta_j\equiv\xi_j\eta_n\equiv 0\pmod p$. Hence we have $\eta_j\equiv\xi_j\eta_n\overline{\xi_n} \pmod{p^\rho}$, so we can choose $a\equiv\eta_n\overline{\xi_n}\pmod{p^\rho}$. This proves (a). Now write $y=ax+p^\rho b$ for a suitable $b\in\mathbb{Z}^n$. Then

$$y^\top Ay \equiv a^2x^\top Ax + 2ap^\rho\, x^\top Ab \pmod{p^{2\rho}}$$

and

$$x^\top Ay = ax^\top Ax + p^\rho x^\top Ab.$$

This implies (b), and the proof shows in particular that (3.1) is independent of the choice of the representative of $a \in (\mathbb{Z}/p^\rho\mathbb{Z})^*$.

Lemma 4. *Let $0 < \alpha < 1$. Equip $\mathbb{R}^n$ with an inner product given by a quadratic form Q. Let $A \subseteq \mathbb{R}^n \setminus \{0\}$ be such that the angle between any two elements of A is at least α. Then $|A| \ll_Q \alpha^{-(n-1)}$.*

Proof. There exists a constant $c > 0$ depending on the choice of the inner product such that the Euclidean angle between any two elements of A is at least $c\alpha$. By appropriate scaling we may assume that each vector in A is on some face of the unit cube $[-1,1]^n \subset \mathbb{R}^n$ (i.e. its sup-norm is one). Now divide each face of this cube into $(n-1)$-dimensional cubes of side-length $< c\alpha/(2n)$. Clearly there are at most $O(\alpha^{-(n-1)})$ such small cubes, and each of them intersects at most one vector from A.

Lemma 5. *Let $m, r \in \mathbb{N}$ and $A \geq 2$. Let $K \subseteq \mathbb{R}$ be a real number field and let $\bar{K} \subseteq \mathbb{C}$ be its Galois closure. For $1 \leq j \leq r$ let $b_j = (b_{1j}, \dots, b_{mj})^\top \in \mathbb{R}^m$ and assume that all b_{ij} are in the ring of integers $\mathcal{O}_K$ and satisfy $b_{ij} = 0$ or*

$$A^{-1} \leq |\sigma(b_{ij})| \leq A \tag{3.2}$$

for all $\sigma \in \mathrm{Gal}(\bar{K}/\mathbb{Q})$. Let $H = \bigcap_j b_j^\perp$. Then the following holds:

(a) We have $\mathrm{dist}(v, H) \ll A^{O(1)} \max_j |\langle b_j, v\rangle|$ for all $v \in \mathbb{R}^m$.

(b) If $H \neq \{0\}$, there is an $\mathbb{R}$-basis $\{v_i\}$ of H with entries in $\mathcal{O}_K$ and $\|v_i\| \ll A^{O(1)}$.

Here all implied constants depend at most on m, r and $\deg(\bar{K}/\mathbb{Q})$.

Proof. For a fixed number $\alpha \geq 1$, we say that an element of K is α-well-balanced, if it can be written as a fraction a/b with $a, b \in \mathcal{O}_K$ and either $a = 0$ and $b = 1$ or

$$A^{-\alpha} \leq |\sigma a|,\ |\sigma b| \leq A^\alpha$$

for each $\sigma \in \mathrm{Gal}(\bar{K}/\mathbb{Q})$. Obviously if a/b is α-well-balanced, then so is $-a/b$, and if in addition $a/b \neq 0$, then also b/a is α-well-balanced. If a/b and c/d are both α-well-balanced, then obviously their product ac/bd is 2α-well-balanced. Finally we claim that also $a/b + c/d$ is β-well-balanced where $\beta = (2\alpha+1)\deg(\bar{K}/\mathbb{Q})$. Indeed, if the sum is zero, then we are done. Otherwise write it as $(ad+bc)/bd$. Clearly $A^{-2\alpha} \leq |\sigma(bd)| \leq A^{2\alpha}$ and $|\sigma(ad+bc)| \leq |\sigma(ad)| + |\sigma(bc)| \leq 2A^{2\alpha} \leq A^{2\alpha+1}$ for each $\sigma \in \mathrm{Gal}(\bar{K}/\mathbb{Q})$. On the other hand,

$$\prod_\sigma |\sigma(ad+bc)| = |\mathcal{N}(ad+bc)| \geq 1$$

so that together with the upper bound we obtain the desired lower bound $A^{-(2\alpha+1)\deg(\bar{K}/\mathbb{Q})} \leq |\sigma(ad+bc)|$ for each $\sigma \in \mathrm{Gal}(\bar{K}/\mathbb{Q})$.

Now we prove part (a). Take a maximal set of independent row vectors $u_1^T, \dots, u_{m'}^T$ of $b_1^T, \dots, b_r^T$ (i.e. $\dim H = m - m'$). Then $u_1, \dots, u_{m'}$ is a basis in $H^\perp$. Following Gram-Schmidt, we obtain inductively an ortho*gonal* basis $u_j' := u_j - \sum_{i<j} u_i' \langle u_j, u_i'\rangle / \|u_i'\|^2$ with entries in K. Then

$$\mathrm{proj}_{H^\perp} v = \sum_{j=1}^{m'} \frac{\langle v, u_j'\rangle}{\langle u_j', u_j'\rangle} u_j'.$$

Each entry in each u_j is $O(1)$-well-balanced, which then implies the same for u_j'. Also by linearity, $\langle v, u_j'\rangle$ is a linear combination of $\langle v, u_j\rangle$'s with $O(1)$-well-balanced coefficients. From this, the statement is obvious.

Now we prove part (b). Let C be a matrix composed of a maximal number of independent rows $b_1^T, \dots, b_r^T$. Its rank is $m' < m$, so there is a nonsingular $m' \times m'$ submatrix in C. By changing the coordinates, we may assume that C is of the block form $(C_1|C_2)$ where C_1 is an invertible $m' \times m'$ matrix and C_2 is an $m' \times (m - m')$ matrix. Hence any vector $y \in H$ can be decomposed as $y = (y_1, y_2) \in \mathbb{R}^{m'} \times \mathbb{R}^{m-m'}$ with $y_1 = -C_1^{-1}C_2 y_2$. Since the entries of $b_1, \dots, b_r$ are $O(1)$-well-balanced, the same holds for $-C_1^{-1}C_2$.

Letting y_2 run through the standard basis in the last $m - m' > 0$ coordinates (since $H \neq \{0\}$), we obtain a basis of H of vectors in K^m with denominators bounded by $A^{O(1)}$. This gives easily the claim.

4. Counting matrices

We start by fixing some **notation and conventions** valid for the rest of this paper. Let Sym_n be the vector space of all symmetric n-by-n matrices, equipped with the standard basis, and let Pos_n be the subset of positive definite matrices. Fix a non-empty open bounded subset $\mathcal{M} \subseteq \overline{\mathcal{M}} \subseteq \mathrm{Pos}_n$ (where the bar denotes the topological closure), and another non-empty open bounded set $\mathcal{M}^*$ whose closure is contained in $\mathcal{M}$. For $Q \in \mathcal{M}$ we obtain an inner product $\langle ., . \rangle_Q$ and a corresponding norm $\|.\|_Q$.

For a rational matrix $Q \in \mathrm{Mat}_n(\mathbb{Q})$ we denote by $\mathrm{den}(Q)$ the smallest positive integer r such that rQ is integral. Let $Q \in \mathrm{Mat}_n(\mathbb{Q})$ be a symmetric positive-definite rational matrix and $\tilde{Q} = \mathrm{den}(Q) \cdot Q = (\tilde{Q}_{ij}) \in \mathrm{Mat}_n(\mathbb{Z})$. Let $\mathcal{Q} = \{\tilde{Q}_{jj} \mid 1 \leq j \leq n\}$ be the set of diagonal entries of $\tilde{Q}$ and $\mathcal{D} = \{\tilde{Q}_{ii}\tilde{Q}_{jj} - \tilde{Q}_{ij}^2 \mid 1 \leq i < j \leq n\} \subseteq \mathbb{N}$ be the set of all 2-by-2 diagonal determinants. We say that a prime p is *Q-good* if p is coprime to all elements in $\mathcal{Q}$ and $-d$ is a quadratic non-residue modulo p for each $d \in \mathcal{D}$.

In the following all implied constants may depend on $\mathcal{M}$ and $\mathcal{M}^*$ (and hence on n) as well as on ε wherever applicable. In particular, for any $Q \in \mathcal{M} \subseteq \overline{\mathcal{M}}$ we can apply Lemma 2 and Lemma 4 with an implied constant that by compactness depends only on $\overline{\mathcal{M}}$, but not on Q. All constants $c_1, c_2, \ldots$ are chosen sufficiently large and may depend on n, but on nothing else.

We return to the problem of estimating $\mathcal{S}(Q, a, b, M)$ defined in (2.7) and provide two bounds in special situations that refer to the Cases 3 and 2, respectively, in the introduction. Note that any $\gamma \in \mathcal{S}(Q, a, b, M)$ satisfies

$$\|\gamma\| \ll (ab^{n-1})^{1/n} \tag{4.1}$$

for $Q \in \mathcal{M}$. Moreover, since $\Delta_1(\gamma) = 1$, each $\gamma \in \mathcal{S}(Q, a, b, M)$ has a column that is not completely divisible by any given prime. We will always assume without loss of generality that this is the first column. We generally write $Q = (Q_{ij})$ and $\gamma = (\gamma_{ij})$ as well as $\gamma_j = (\gamma_{1j}, \ldots, \gamma_{nj})^\top$ for the j-th column of $\gamma \in \mathcal{S}(Q, a, b, M)$. Then

$$\gamma_i^\top Q \gamma_j = (ab^{n-1})^{2/n} Q_{ij} + O\left((ab^{n-1})^{(2-M)/n}\right) \tag{4.2}$$

for $1 \leq i, j \leq n$, and $\Delta_2(\gamma) = b$ implies

$$\gamma_{ij}\gamma_{i'j'} - \gamma_{i'j}\gamma_{ij'} \equiv 0 \,(\mathrm{mod}\ b) \tag{4.3}$$

for all $1 \leq i, i', j, j' \leq n$. The following lemma makes precise the argument of Case 3 in the introduction. It feature the same factor $p^{\nu(n-2)+\varepsilon}$ and is uniform in the height of the quadratic form. A key ingredient is Lemma 3.

Lemma 6. *Let $Q \in \mathcal{M} \cap \mathrm{Mat}_n(\mathbb{Q})$, $1 \leq \nu \leq n$, $p > 2$ a prime. Then*

$$|\mathcal{S}(Q, p^\nu, p^\nu, \infty)| \ll \mathrm{den}(Q)^{\frac{1}{2}(n-1)^2} p^{\nu(n-2+\varepsilon)}$$

for any $\varepsilon > 0$.

Proof. Without loss of generality assume that $p \nmid \gamma_{11}$. By (4.2) with $i = j = 1$ and Lemma 2 with $m = p^\nu$ we can choose the first column γ_1 of γ in $p^{\nu(n-2+\varepsilon)}$ ways.

Fix $\mu \geq 0$. We count the number of choices for the second column $\gamma_2 = (\gamma_{12}, \ldots, \gamma_{n2})^\top$ of γ such that

$$\min_i v_p(\gamma_{i2}) = \mu \tag{4.4}$$

where v_p denotes the p-adic valuation. By (4.1) we clearly have $\mu = O(1)$, and in fact if $\mu \geq \nu$, there are at most $O(1)$ choices for each γ_{i2} by (4.1), hence $O(1)$ choices for γ_2. Now let $\mu < \nu$. Let $\tilde{Q} := \mathrm{den}(Q) \cdot Q =$

$(\tilde{Q}_{ij}) \in \mathrm{Mat}_n(\mathbb{Z})$. Let $x = (\xi_1, \dots, \xi_n)^\top, y = (\eta_1, \dots, \eta_n)^\top \in \mathbb{Z}^n$ be two choices for γ_2 satisfying (4.4). By (4.2) with $i = j = 2$ and the definition of the set $\mathcal{S}(Q, p^\nu, p^\nu, \infty)$ we have

$$\|x\|_{\tilde{Q}}^2 = \|y\|_{\tilde{Q}}^2 = p^{2\nu}\tilde{Q}_{22} \equiv 0 \,(\mathrm{mod}\ p^{2\nu}). \tag{4.5}$$

On the other hand, by (4.3) we have $p^{-\mu}(\gamma_{i1}\xi_{i'} - \gamma_{i'1}\xi_i) \equiv 0 \,(\mathrm{mod}\ p^{\nu-\mu})$. Lemma 3(a) implies $p^{-\mu}x \equiv a_1\gamma_1 \,(p^{\nu-\mu})$ and similarly $p^{-\mu}y \equiv a_2\gamma_1 \,(\mathrm{mod}\ p^{\nu-\mu})$ for some $a_1, a_2 \in (\mathbb{Z}/p^{\nu-\mu}\mathbb{Z})^*$, hence

$$p^{-\mu}x \equiv p^{-\mu}a_1\bar{a}_2 y \,(\mathrm{mod}\ p^{\nu-\mu}),$$

which in turn implies $p^{-2\mu}(\xi_i\eta_j - \xi_j\eta_i) \equiv 0 \,(\mathrm{mod}\ p^{\nu-\mu})$. By Lemma 3(b) in connection with (4.5), we conclude $\langle p^{-\mu}x, p^{-\mu}y\rangle_{\tilde{Q}} \equiv 0 \,(\mathrm{mod}\ p^{2\nu-2\mu})$ or

$$\langle x, y\rangle_{\tilde{Q}} \equiv 0 \,(\mathrm{mod}\ p^{2\nu})$$

for $p > 2$. Hence either x and y are collinear, or the angle between x and y (with respect to $\langle ., . \rangle_{\tilde{Q}}$ which determines the same angles as the inner product $\langle ., . \rangle_Q$) is

$$\geq \arccos\frac{\tilde{Q}_{22} - 1}{\tilde{Q}_{22}} \gg \frac{1}{\tilde{Q}_{22}^{1/2}} \gg \frac{1}{\mathrm{den}(Q)^{1/2}}.$$

By Lemma 4 there are $\ll \mathrm{den}(Q)^{(n-1)/2}$ choices for γ_2. The same argument applies for all other columns, and the proof is complete.

The following lemma makes precise the argument of Case 2 in the introduction. The main point assumptions here is p and q are Q-good primes, which is the analogue of $p, q \equiv 3 \,(\mathrm{mod}\ 4)$ in the simplified version in the introduction. Again Lemma 3 is a key ingredient.

Lemma 7. *Let $\nu \in \{n, n/2\} \cap \mathbb{Z}$. Then for every $Q \in \mathrm{Mat}_n(\mathbb{Q}) \cap \mathcal{M}$ and every two different Q-good primes p and q the estimate*

$$|\mathcal{S}(Q, q^\nu, p^\nu, \infty)| \ll \left(1 + \frac{q}{p}\right)^{\nu(n-1)} \left(q^{1/n}p^{(n-1)/n}\right)^{\nu(n-2+\varepsilon)}$$

holds.

Proof. Write $2\nu/n = \kappa \in \{1, 2\}$. Assume without loss of generality that the first column γ_1 of γ is not completely divisible by p. As in the proof of Lemma 6 we conclude from Lemma 2 with $m = (qp^{n-1})^{\nu/n}$ that there are

$$\ll (q^{1/n}p^{(n-1)/n})^{\nu(n-2+\varepsilon)}$$

ways to choose γ_1. If all other columns are completely divisible by p^ν, then by (4.1) there are $O(1+(q/p)^{\nu/n})$ choices for each entry of $\gamma_2, \dots, \gamma_n$. This is admissible. Otherwise assume without loss of generality that the second column γ_2 of γ satisfies (4.4) with $0 \leq \mu < \nu$. Write as before $\mathrm{den}(Q) \cdot Q = (\tilde{Q}_{ij}) \in \mathrm{Mat}_n(\mathbb{Z})$. By Lemma 3(b) with $\rho = \nu - \mu = \kappa n/2 - \mu$ and (4.2) with $(i, j) \in \{(1, 2), (1, 1), (2, 2)\}$ and $M = \infty$ we conclude that

$$2p^{-\mu}\tilde{Q}_{12}q^\kappa p^{\kappa(n-1)} \equiv a\tilde{Q}_{11}q^\kappa p^{\kappa(n-1)} + \bar{a}p^{-2\mu}\tilde{Q}_{22}q^\kappa p^{\kappa(n-1)} \,(\mathrm{mod}\ p^{\kappa n-2\mu})$$

for some $a \in (\mathbb{Z}/p^{n-\mu}\mathbb{Z})^*$, i.e. $a^2p^{2\mu}\tilde{Q}_{11} - 2ap^\mu\tilde{Q}_{12} + \tilde{Q}_{22} \equiv 0 \,(\mathrm{mod}\ p^\kappa)$ and a fortiori

$$a^2p^{2\mu}\tilde{Q}_{11} - 2ap^\mu\tilde{Q}_{12} + \tilde{Q}_{22} \equiv 0 \,(\mathrm{mod}\ p).$$

The case $\mu > 0$ leads immediately to a contradiction since $p \nmid \tilde{Q}_{22}$ by definition of Q-goodness. In the case $\mu = 0$, we see that $\tilde{Q}_{12}^2 - \tilde{Q}_{11}\tilde{Q}_{22}$ must be a quadratic residue modulo p which again contradicts that p is Q-good.

5. The exchange lemma

For $\gamma \in \mathrm{Mat}_n(\mathbb{Z})$ and $m \in \mathbb{N}$ define the linear map

$$B_{\gamma,m} : \mathrm{Sym}_n \to \mathrm{Sym}_n, \quad Q \mapsto \gamma^\top Q\gamma - m^{1/n}Q.$$

The following crucial lemma enables us to "exchange" the matrix Q in $\mathcal{S}(Q, a, b, M)$ for a matrix Q' that has better diophantine properties. As mentioned in the introduction, we will ue this to exchange Q with a matrix that has better diophantine properties.

Lemma 8. *There exist constants c_1, c_2 with the following property.*

Let $L > 2$, $D \geq 1$, $M \geq c_1 D$. Let $I := [L, 2L^D]$ and let $\mathcal{P} \subseteq \{(p^\nu, q^\nu) \mid p, q \in I, 1 \leq \nu \leq n\}$ be a set of pairs of prime powers. Let $Q \in \mathcal{M}^*$. Then there exists a subspace $\{0\} \neq H \subseteq \mathrm{Sym}_n$ (depending on Q, $\mathcal{P}$ and M) defined in (5.4) below, such that for every every matrix $Q' \in H \cap \mathcal{M}$ one has

$$\mathcal{S}(Q, q^\nu, p^\nu, M) \subseteq \mathcal{S}(Q', q^\nu, p^\nu, \infty) \quad \text{for all } (p^\nu, q^\nu) \in \mathcal{P}. \tag{5.1}$$

Moreover, there exists a subset $\mathcal{P}' \subseteq \mathcal{P}$ with $|\mathcal{P}'| \leq n(n+1)/2$ such that, setting

$$K := \mathbb{Q}\big((qp^{n-1})^{2\nu/n} : (p^\nu, q^\nu) \in \mathcal{P}'\big), \tag{5.2}$$

there exists a matrix $Q' \in H \cap \mathcal{M} \cap \mathrm{Mat}_n(K)$, and if $K = \mathbb{Q}$, then

$$\mathrm{den}(Q') \ll L^{c_2 D}. \tag{5.3}$$

Proof. Define[3]

$$H := \bigcap_{\substack{(p^\nu,q^\nu)\in\mathcal{P}\\ \gamma\in\mathcal{S}(Q,q^\nu,p^\nu,M)}} \ker B_{\gamma,q^{2\nu}p^{2\nu(n-1)}}. \tag{5.4}$$

Then by definition, (5.1) is satisfied for all $Q' \in H \cap \mathcal{M}$. To each $B_{\gamma,q^{2\nu}p^{2\nu(n-1)}}$ we can associate a matrix. Take a minimal set of rows $b_1^\top, \dots, b_r^\top \in \mathbb{R}^n$, $r \leq n(n+1)/2$, of these matrices that generate $H^\perp$. Let $\mathcal{P}'$ be the set of corresponding pairs (p^ν, q^ν), and define K as in (5.2). Then the b_j have entries that are in $\mathbb{Z}$ or of the form $a - (qp^{n-1})^{2\nu/n}$ with $(p^\nu, q^\nu) \in \mathcal{P}'$ and $a \in \mathbb{Z}$. In particular, they are either 0, or by the considerations in the beginning of the proof of Lemma 5 they satisfy (3.2) with $A \ll L^{c_3 D}$ for some $c_3 > 0$ (with an implied constant depending only on n). By Lemma 5(a) we have $\mathrm{dist}(Q, H) \ll L^{c_4 D - M}$ for a constant c_4. Hence for $M \geq c_4 D + c_5$, the subspace H intersects $\mathcal{M}$ in a ball of fixed radius (recall that $Q \in \mathcal{M}^* \subseteq \overline{\mathcal{M}^*} \subseteq \mathcal{M}$). In particular, $H = \{0\}$ is impossible. It follows now from Lemma 5(b) that we can choose $Q' \in H \cap \mathcal{M} \cap \mathrm{Mat}_n(K)$ that in the case $K = \mathbb{Q}$ satisfies (5.3).

6. A recursive argument

We are now ready to prove good upper bounds for $\mathcal{S}(Q, p^\nu, q^\nu, M)$ for suitable primes in suitable ranges.

Let $Q \in \mathcal{M}^*$ and $L > 2$. Let $M, D_1, D_2 \geq 1$ be (large, but fixed) parameters satisfying

$$M \geq c_1 D_1^{n(n+1)/2} D_2^{n(n+1)/2+1}. \tag{6.1}$$

For $0 \leq j \leq n(n+1)/2$ let

$$I_j := [L, 2L^{D_1^j D_2^{j+1}}], \quad \mathcal{P}_j = \{(p^\nu, q^\nu) \mid p, q \in I_j, 1 \leq \nu \leq n\},$$

and with this choice of I_j and $\mathcal{P}_j$ let $H_j \subseteq \mathrm{Sym}_n$ be as in (5.4). Attached to these data is a field K_j and a matrix $Q_j \in \mathcal{M} \cap \mathrm{Mat}_n(K_j) \cap H_j$ as in Lemma 8. Clearly $\mathrm{Sym}_n \supseteq H_0 \supseteq H_1 \supseteq \dots$. Therefore we must have $H_i = H_{i+1}$ for some $i < n(n+1)/2$. Fix once and for all such an index i. Since $Q_i \in H_i = H_{i+1}$, it follows from (5.1) that

$$\mathcal{S}(Q, q^\nu, p^\nu, M) \subseteq \mathcal{S}(Q_i, q^\nu, p^\nu, \infty) \quad \text{for all } (p^\nu, q^\nu) \in \mathcal{P}_{i+1}.$$

[3]The empty intersection is just Sym_n.

Write $Q_i = (Q_{rs})_{1\le r,s\le n} \in \mathrm{Mat}_n(K_i)$ and choose any (r,s) with $Q_{rs}\neq 0$. Then by (4.2) any $\gamma \in \mathcal{S}(Q_i, q^\nu, p^\nu, \infty)$ satisfies

$$K_i \ni Q_{rs}^{-1}\gamma_r^\top Q\gamma_s = (qp^{n-1})^{2\nu/n}.$$

Recall that K_i is contained in a finite extension of $\mathbb{Q}$ by n-th roots of primes in I_i. In particular, if $p\neq q\in I_{i+1}\setminus I_i$ and $2\nu/n\notin\mathbb{N}$, then the right hand side is not in K_i (see e.g. [Be]), a contradiction. This is the same contradiction as in Case 1 in the introduction. We conclude

$$|\mathcal{S}(Q,q^\nu,p^\nu,M)|\le |\mathcal{S}(Q_i,q^\nu,p^\nu,\infty)| = 0,\quad p\neq q\in I_{i+1}\setminus I_i,\ \frac{2\nu}{n}\notin\mathbb{N}. \tag{6.2}$$

Let us now consider the cases (i) $q=p$, or (ii) $q\neq p$, but $2\nu/n\in\mathbb{N}$; both cases together are equivalent to $(qp^{n-1})^{2\nu/n}\in\mathbb{N}$. We run a similar, but slightly more complicated argument. Let $\mathcal{L} := L^{(D_1D_2)^{i+1}}$ and for $0\le j\le n(n+1)/2$ define

$$I_j^* = [\mathcal{L}, 2\mathcal{L}^{D_1^j}],\quad \tilde{I}_j^* = [\mathcal{L}^{D_1^j}, 2\mathcal{L}^{D_1^j}].$$

If we assume that

$$D_2 \ge D_1^{n(n+1)/2} \tag{6.3}$$

then $I_0^*\subseteq I_1^*\subseteq\ldots\subseteq I_{n(n+1)/2}^*\subseteq I_{i+1}\setminus I_i$. We attach inductively to each interval I_j^* a subspace $H_j^*\subseteq \mathrm{Sym}_n$, a matrix $Q_j^*\in\mathcal{M}\cap\mathrm{Mat}_n(\mathbb{Q})\cap H_j^*$ and a set $\mathcal{P}_j^*$ of pairs of prime powers as follows: let

$$\mathcal{P}_0^* := \{(p^\nu,q^\nu)\mid p,q\in I_0^*, 1\le\nu\le n, (qp^{n-1})^{2\nu/n}\in\mathbb{N}\},$$

and for $j>0$ let

$$\begin{aligned}\tilde{\mathcal{P}}_j^* &:= \{(p^\nu,q^\nu)\mid p,q\in\tilde{I}_j^*, 1\le\nu\le n, (qp^{n-1})^{2\nu/n}\in\mathbb{N}, p,q \text{ are } Q_{j-1}^*\text{-good}\},\\ \mathcal{P}_j^* &:= \mathcal{P}_{j-1}^*\cup\tilde{\mathcal{P}}_j^*.\end{aligned}$$

With this choice of I_j^* and $\mathcal{P}_j^*$ let H_j^* be as in (5.4). Note that in our present situation, the number field (5.2) is always $\mathbb{Q}$. Let $Q_j^*\in\mathcal{M}\cap\mathrm{Mat}_n(\mathbb{Q})\cap H_j^*$ be as in Lemma 8 satisfying (5.3). Clearly, $\mathrm{Sym}_n\supseteq H_0^*\supseteq H_1^*\supseteq\ldots$. Therefore we must have $H_k^*=H_{k+1}^*$ for some $k<n(n+1)/2$. Fix once and for all such an index k. Since $Q_k^*\in H_k^*=H_{k+1}^*$, it follows from (5.1) that $\mathcal{S}(Q,q^\nu,p^\nu,M)\subseteq\mathcal{S}(Q_k^*,q^\nu,p^\nu,\infty)$ for all $(p^\nu,q^\nu)\in\mathcal{P}_{k+1}^*$ and hence a fortiori for all $(p^\nu,q^\nu)\in\tilde{\mathcal{P}}_{k+1}^*$. Recalling that the latter set consists of powers of Q_k^*-good primes, we conclude from Lemma 7 that

$$|\mathcal{S}(Q,q^\nu,p^\nu,M)|\le|\mathcal{S}(Q_k^*,q^\nu,p^\nu,\infty)|\ll p^{\nu(n-2+\varepsilon)},\quad (p^\nu,q^\nu)\in\tilde{\mathcal{P}}_{k+1}^*,\quad p\neq q, \tag{6.4}$$

(here we use that $p^\nu\asymp q^\nu$ for $(p^\nu,q^\nu)\in\tilde{\mathcal{P}}_{k+1}^*$), and from Lemma 6 and (5.3) that

$$|\mathcal{S}(Q,p^\nu,p^\nu,M)|\le|\mathcal{S}(Q_k^*,p^\nu,p^\nu,\infty)|\ll\mathcal{L}^{c_6D_1^k}p^{\nu(n-2+\varepsilon)},\quad (p^\nu,p^\nu)\in\tilde{\mathcal{P}}_{k+1}^*. \tag{6.5}$$

We recall that $p\ge\mathcal{L}^{D_1^{k+1}}$ for $p\in\tilde{\mathcal{P}}_{k+1}^*$. Combining (6.2), (6.4), (6.5), we obtain the following central result which concludes our diophantine investigations.

Proposition 1. *Let $L>2$ and let $M,D_1,D_2\ge1$ be satisfying (6.3) and (6.1). Let $Q\in\mathcal{M}^*$. There exist $0\le i,k<n(n+1)/2$ and two sets $\mathcal{D},\mathcal{Q}\subseteq\mathbb{N}$ (depending on Q) of cardinality at most $n(n-1)/2$ and n, respectively, with the following properties.*

Put $\mathcal{L}:=L^{(D_1D_2)^{i+1}}$. Then it holds that $\mathcal{D},\mathcal{Q}\subseteq[1,O(\mathcal{L}^{c_7D_1^k})]$. Let $\mathcal{P}$ be the set of all primes p in $[\mathcal{L}^{D_1^{k+1}},2\mathcal{L}^{D_1^{k+1}}]$ coprime to all elements in $\mathcal{Q}$ and such that $-d$ is a quadratic non-residue modulo p for each $d\in\mathcal{D}$. Then

$$|\mathcal{S}(Q,q^\nu,p^\nu,M)|\ll p^{\nu(n-2+\varepsilon)+\frac{c_6}{D_1}}$$

for all $p,q\in\mathcal{P}$ and $1\le\nu\le n$.

7. Completion of the proof of Theorem 1

In order to use Proposition 1, we need to make sure that sufficiently many primes satisfy the conditions of the proposition, in other words $\mathcal{P}$ is sufficiently large and in particular non-empty. To this end we use the following Linnik-type result. Let as usual $\Lambda(n)$ denote the von Mangoldt function.

Lemma 9. *There exists an absolute constant $c > 0$ such that*

$$\sum_{\substack{x\le n\le 2x\\ n\equiv a \,(\mathrm{mod}\ m)}} \Lambda(n) \gg \frac{x}{m^{3/2}}$$

for all integers $a \in \mathbb{Z}$, $m \geq 2$ with $(a, m) = 1$, provided $x \geq m^c$.

Proof. This is [IK, Corollary 18.8] for the summation condition $n \leq x$, and the proof for a dyadic interval is essentially identical, starting from the asymptotic formula in [IK, Proposition 18.5] (see also [GS, Corollary on p.572]) .

By the Chinese remainder theorem and quadratic reciprocity, the set $\mathcal{P}$ in Proposition 1 can be described by congruence conditions modulo a number $m \ll \mathcal{L}^{c_8 D_1^k}$ for some c_8, and we impose in addition that all elements in $\mathcal{P}$ are coprime to the level N of Γ. We conclude from Lemma 9 that there exists a constant c_9 such that

$$D_1 \geq c_9 \geq 3c_6$$

implies that

$$|\mathcal{P}| \gg \mathcal{L}^{D_1^{k+1} - c_{10} D_1^k} \geq \mathcal{L}^{\frac{1}{2} D_1^{k+1}},$$

provided right hand side exceeds $10 \log N$, a generous multiple of the number of distinct prime factors of N. With this choice of D_1 we now specify the other parameters in Proposition 1. We fix some D_2 and M satisfying (6.3) and (6.1), and we put

$$L = 10(\log N) D(\mu^*)^\eta$$

for some small constant $\eta > 0$ to be specified in a moment and define $\mathcal{L} = L^{(D_1 D_2)^{i+1}}$ as in Proposition 1. Finally we choose $\mathcal{M}^*$ to contain the image of Ω under the map $gK \mapsto (\det g)^2 g^{-\top} g^{-1}$, cf. (2.6), so that Proposition 1 is applicable for the matrix Q in question. Now we return to (2.8), which we apply with

$$L_0 = \mathcal{L}^{D_1^{k+1}} = L^{(D_1 D_2)^{i+1} D_1^{k+1}} = \Big(10(\log N) D(\mu^*)^\eta\Big)^{(D_1 D_2)^{i+1} D_1^{k+1}}.$$

This gives

$$|F(g)|^2 \ll N^\varepsilon D(\mu^*) \left(\frac{1}{L_0^{1/2}} + D(\mu^*)^{-\frac{1}{n(n-1)} + \eta(n^3 + \frac{M}{2})(D_1 D_2)^{i+1} D_1^{k+1}} + \frac{1}{L_0^{1/2}} \right),$$

where $i, k < n(n+1)/2$ and M satisfies (6.1). Choosing $\eta = \eta(n) > 0$ sufficiently small, it is clear that we can obtain (1.3) for some $\delta > 0$.

References

[Be] A. Besicovitch, *On the linear independence of fractional powers of integers*, J. London. Math. Soc. **15** (1940), 3-6

[BH] V. Blomer, G. Harcos, *Twisted L-functions over number fields and Hilberts eleventh problem*, Geom. Funct. Anal. **20** (2010), 1-52; erratum available at the authors webpages

[BHM] V. Blomer, G. Harcos, D. Milićević, *Eigenfunctions on arithmetic hyperbolic 3-manifolds*, to appear in Duke Math. J.

[BM] V. Blomer, P. Maga, *The sup-norm problem for* PGL(4), IMRN **2015** (vol. 14), 5311-5332

[BP] V. Blomer, A. Pohl, *The sup-norm problem on the Siegel modular space of rank 2*, to appear in Amer. J. Math.

[BT] F. Brumley and N. Templier, *Large values of cusp forms on* GL(n), `arXiv:1411.4317`

[Ca] J. Cassels, *Rational quadratic forms*, L.M.S. Monographs, No.13, 1978.

[GRS] A. Ghosh, A. Reznikov, P. Sarnak, *Nodal domains of Maass forms I*, Geom. Funct. Anal. **23** (2013), 1515-1568.

[GS] D. Goldfeld, A. Schinzel, *On Siegel's zero*, Ann. Scuola Norm. Sup. Pisa **4** (1975), 571-583

[HM] G. Harcos, P. Michel, *The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points. II*, Invent. Math. **163** (2006), 581-655

[HRR] R. Holowinsky, G. Ricotta, E. Royer, *On the sup-norm of* SL_3 *Hecke-Maass cusp forms*, arXiv:1404.3622
[IK] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, AMS Colloquium publications **53**, Providence 2004
[IS] H. Iwaniec, P. Sarnak, L^∞ *norms of eigenfunctions of arithmetic surfaces*, Ann. Math. **141** (1995), 301-320.
[Mag] P. Maga, *Shifted convolution sums and Burgess type subconvexity over number fields*, arXiv:1312.0553
[Mar] S. Marshall, *Sup norms of Maass forms on semisimple groups*, arXiv:1405.7033
[Ru] Z. Rudnick, *On the asymptotic distribution of zeros of modular forms*, Int. Math. Res. Not. 2005, no. 34, 2059-2074
[RS] Z. Rudnick, P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), 195-213
[Sa1] P. Sarnak, *Letter to Morawetz*, available at http://www.math.princeton.edu/sarnak
[Sa2] P. Sarnak, *Recent progress on the quantum unique ergodicity conjecture*, Bull. Am. Math. Soc. **48** (2011), 211-228
[Ze] S. Zelditch, *Recent developments in mathematical quantum chaos*, Current developments in mathematics, 2009, Somerville, MA: International Press, 2010, pp. 115-204.

Mathematisches Institut, Bunsenstr. 3-5, 37073 Göttingen, Germany
E-mail address: vblomer@math.uni-goettingen.de

MTA Alfréd Rényi Institute of Mathematics, POB 127, Budapest H-1364, Hungary
E-mail address: maga.peter@renyi.mta.hu