

SUMS OF SQUARES IN REAL QUADRATIC FIELDS AND HILBERT MODULAR GROUPS

FERNANDO CHAMIZO AND ROBERTO J. MIATELLO

ABSTRACT. We use the spectral theory of Hilbert-Maass forms for real quadratic fields to obtain the asymptotics of some sums involving the number of representations as a sum of two squares in the ring of integers.

1. INTRODUCTION

The asymptotic study of the average of the arithmetical function $r(n)$ giving the number of representations of n as a sum of two squares is the goal of the celebrated Gauss circle problem. It asks for the infimum of the exponents α 's satisfying

$$(1.1) \quad \sum_{n \leq x} r(n) = \pi x + O(x^\alpha).$$

The left-hand side counts the number of lattice points in a circle of radius \sqrt{x} and πx gives its area. Gauss used an approximation of this kind when studying the class number of quadratic forms [Gau81]. A simple geometric reasoning, already employed by Gauss, shows (1.1) for $\alpha = 1/2$. Sierpiński [Sie74] proved the estimate for $\alpha = 1/3$ with a rather complicated argument (see [Hux96] for a short modern elementary approach). On the other hand, $\alpha = 1/3$ can be obtained in a more direct way using the Euclidean spectral expansion (i.e., classical Fourier analysis) of radial functions on \mathbb{R}^2 [Lan93, VIII§8].

A finer asymptotic property to be studied about $r(n)$ is its self-correlation. In [Iwa02] it is proved that

$$(1.2) \quad \sum_{n \leq x} r(n)r(n+1) = 8x + O(x^{2/3})$$

and there are similar formulas replacing $r(n+1)$ by $r(n+k)$ (see [Cha99] for the uniformity). Although (1.1) for $\alpha = 1/3$ and (1.2) seem unrelated, from the analytic point of view one can run both proofs along similar lines changing Euclidean spectral expansions based on Fourier series by hyperbolic spectral expansions based on Maass forms and Eisenstein series. This second situation is by far more involved and one has to bypass unsettled problems like the existence of exceptional eigenvalues or the L^∞ bounds of the eigenfunctions, which become trivial in the Euclidean setting. It is noteworthy to mention that the error term in (1.2) remains unimproved

2010 *Mathematics Subject Classification.* 11F72, 11F41, 11N45.

The first named author is partially supported by the MTM2017-83496-P grant of the MICINN (Spain) and by "Severo Ochoa Programme for Centres of Excellence in R&D" (SEV-2015-0554).

while van der Corput method and other finer techniques of exponential sums have proved (1.1) for some $\alpha < 1/3$ [Hux96], namely the best known result [BW17] allows to take α slightly smaller than 0.314.

Several authors have considered the analogue of Gauss circle problem in totally real number fields [Sie36a], [Sch65], [Sch62], [Rau88] where $n \in \mathbb{Z}$ is replaced by $\lambda \in \mathcal{O}$, with \mathcal{O} the ring of integers. A basic issue is to find a real number field analogue of $n \leq x$ (note that the existence of infinitely many units prevents us from using the norm). Following Siegel [Sie36a], a natural condition is to limit the size of every Galois conjugate. In the real quadratic case of discriminant Δ , the main result in [Sch62] implies that

$$(1.3) \quad \sum_{\substack{0 \leq \lambda < V_1 \\ 0 \leq \lambda^\sigma < V_2}} r(\lambda) = \frac{\pi^2}{\Delta} V_1 V_2 + O((V_1 V_2)^\alpha) \quad \text{for any } \alpha > \frac{2}{3}$$

where λ^σ is the real conjugate of λ and $r(\lambda)$ is defined in the natural way:

$$(1.4) \quad r(\lambda) = \#\{(\xi, \eta) \in \mathcal{O}^2 : \lambda = \xi^2 + \eta^2\}.$$

The purpose of this paper is to get an analogue of (1.2) for real quadratic fields by applying the hyperbolic circle problem for products of two upper half-planes. This latter problem was studied in general in [BGM11] for multiple products corresponding to totally real number fields of arbitrary degree. The underlying analysis involves the spectral theory of Hilbert-Maass forms.

To state our main results we use the abbreviations

$$(1.5) \quad N_D(V_1, V_2) = \sum_{\substack{0 \leq \lambda < V_1 \\ 0 \leq \lambda^\sigma < V_2}} r(\lambda)r(\lambda + 1) \quad \text{and} \quad C_D = \frac{32\Delta}{(2 - \chi(2) + 2\chi(4)) \sum_{n=1}^{\Delta} n^2 \chi(n)}$$

where Δ is the discriminant of the quadratic real field $\mathbb{Q}(\sqrt{D})$ we are considering and χ is the character corresponding to the Kronecker symbol $\left(\frac{\Delta}{\cdot}\right)$. See the next section for more details about the notation.

Firstly we state a result like (1.3) for the self-correlation.

Theorem 1.1. *For $V_1, V_2 \rightarrow \infty$*

$$N_D(V_1, V_2) = C_D V_1 V_2 + O((V_1 V_2)^{3/4}).$$

Secondly, we are interested in the global uniformity in V_1 and V_2 that is not considered in [BGM11]. Note that there are limiting arithmetical situations, for instance if V_2 is like V_1^{-1} we are essentially considering solutions of Pell's equation which have an exponential spacing, too sparse to be captured by harmonic analysis and by an asymptotic formula.

Theorem 1.2. *For $0 < V_2 < 1$ and $V_1 V_2^2 \rightarrow \infty$, if there are no exceptional eigenvalues then*

$$N_D(V_1, V_2) = C_D V_1 V_2 + O(V_1^{3/4} V_2^{1/2})$$

and if there exist exceptional eigenvalues then we have to add $O(V_1^{1/2+c} V_2^{1/4})$ with $c = \sup \Im t_{\ell_1}$.

REMARK. In principle one could suspect a chaotic behavior of C_D because of the arithmetic nature of the character sum but it is not difficult to prove that $C_D \asymp D^{-3/2}$. A precise result is included in §7. It may sound surprising that the existence or not of exceptional eigenvalues plays a role in Theorem 1.2 but not in Theorem 1.1. In the terminology in [BGM11], we have a large spectral gap and the influence of the exceptional eigenvalues in Theorem 1.1 is absorbed by the error term. The fundamental result here is the bound of Kim and Shahidi [KS02] for the size of the potential exceptional eigenvalues that allows to take $c = 1/9$ in Theorem 1.2.

2. NOTATION AND BASIC CONCEPTS

We follow mainly the notation of [BM09] and [BGM11]. We recall it briefly reviewing at the same time the basic concepts. We use Ladau's O -notation and Vinogradov's \ll -notation indistinctly.

The Poincaré half-plane \mathbb{H} is the Riemannian manifold given by the upper half-plane $\Im z > 0$ endowed with the hyperbolic metric $y^{-2}(dx^2 + dy^2)$, which induces the invariant measure $d\mu(z) = y^{-2}dx dy$. It is possible to give an explicit formula for the corresponding hyperbolic distance ρ , namely

$$(2.1) \quad \cosh \rho(z, w) = 1 + 2u(z, w) \quad \text{with} \quad u(z, w) = \frac{|z - w|^2}{4\Im z \Im w}.$$

The group $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm \mathrm{Id}\}$ acts faithfully on \mathbb{H} in the standard way by linear fractional transformations and in fact it coincides with the group of orientation preserving isometries of \mathbb{H} . This implies in particular that u is invariant, meaning $u(z, w) = u(\gamma(z), \gamma(w))$ for any $z, w \in \mathbb{H}$ and $\gamma \in \mathrm{PSL}_2(\mathbb{R})$.

For each discrete subgroup $\Gamma < \mathrm{PSL}_2(\mathbb{R})$ such that $\Gamma \backslash \mathbb{H}$ has finite volume, the spectral theory of automorphic forms allows to expand any $f \in L^2(\Gamma \backslash \mathbb{H})$ in terms of the eigenfunctions of the Laplace-Beltrami operator $\Delta = -y^2(\partial_x^2 + \partial_y^2)$. In some sense, the role of the Fourier transform is played in this context by the *Selberg transform*

$$k \in C_0^\infty([0, \infty)) \longmapsto h(t) = \int_{\mathbb{H}} k(u(z, i)) y^{1/2+it} d\mu(z).$$

As in the case of the Fourier transform, we can relax a lot the C_0^∞ regularity still having a sound and useful Selberg transform. It is easier to introduce the conditions in terms of the transform itself, taking for granted the existence of the integral. Following Selberg [Ber54, Satz 3.4], we ask for the existence of a strip $S_\delta = \{z : |\Im z| < 1/2 + \delta\}$ with $\delta > 0$ such that

$$(2.2) \quad h \text{ is holomorphic in } S_\delta \quad \text{and} \quad |h(z)| \ll (|z| + 1)^{-2-\delta} \text{ for } z \in S_\delta.$$

This is satisfied when $k \in C_0^\infty$ (see [Iwa02, §1.8]).

A novelty with respect to the Euclidean setting is that in the cases of arithmetical relevance e.g., $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, there is a discrete spectrum (corresponding mainly to Maass cusp forms) and a continuous spectrum (corresponding to Eisenstein series). This non-classical harmonic

analysis built with nonholomorphic automorphic forms has had a profound impact on analytic number theory specially since the development of Kuznetsov's formula [Haf85].

We focus on the case of real quadratic number fields $\mathbb{Q}(\sqrt{D})$ with $D \in \mathbb{Z}_{>1}$ squarefree. The corresponding ring of integers is

$$\mathcal{O} = \begin{cases} \mathbb{Z}[\sqrt{D}] & \text{if } D \not\equiv 1 \pmod{4}, \\ \mathbb{Z}[(1 + \sqrt{D})/2] & \text{if } D \equiv 1 \pmod{4}. \end{cases}$$

To parallel the previous case $\mathrm{PSL}_2(\mathbb{Z})$, the natural object to work with is the full Hilbert modular group

$$(2.3) \quad \Gamma_{\mathcal{O}} = \{(\gamma, \gamma^\sigma) : \gamma \in \mathrm{PSL}_2(\mathcal{O})\}$$

acting on $\mathbb{H}^2 = \mathbb{H} \times \mathbb{H}$ where γ^σ denotes the action by the nontrivial element in the Galois group of $\mathbb{Q}(\sqrt{D})$ (the real conjugation) on the entries of γ . It turns out that $\Gamma_{\mathcal{O}}$ is a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})^2$ and $\Gamma_{\mathcal{O}} \backslash \mathbb{H}^2$ has finite volume. In general the groups with these two properties are called *lattices* and they are said to be *irreducible* if the projections on each factor of $\mathrm{PSL}_2(\mathbb{R})^2$ are dense. This avoids artificial examples like $\mathrm{PSL}_2(\mathbb{Z})^2$ that can be “reduced” to discrete groups acting on \mathbb{H} .

If Γ is an irreducible lattice, spectral theory allows to analyze $L^2(\Gamma \backslash \mathbb{H}^2)$ in terms of the simultaneous eigenfunctions $\psi = \psi(z_1, z_2)$ of Δ_{z_1} and Δ_{z_2} (where the subscript indicates the variable). Imposing $\psi \in L^2(\Gamma \backslash \mathbb{H}^2)$ we have a discrete sequence of couples of eigenvalues $\{\boldsymbol{\lambda}_\ell\}_\ell$ with $\boldsymbol{\lambda}_\ell = (\lambda_{\ell_1}, \lambda_{\ell_2})$ and corresponding orthonormal eigenfunctions ψ_ℓ ,

$$\Delta_{z_1} \psi_\ell = \lambda_{\ell_1} \psi_\ell, \quad \Delta_{z_2} \psi_\ell = \lambda_{\ell_2} \psi_\ell.$$

We reserve the label $\ell = 0$ for the trivial couple $\boldsymbol{\lambda}_0 = (0, 0)$ and consequently $\psi_0 = |\Gamma \backslash \mathbb{H}^2|^{-1/2}$. It is said that $\boldsymbol{\lambda}_\ell$ is *exceptional* if $0 < \lambda_{\ell_1} < 1/4$ or $0 < \lambda_{\ell_2} < 1/4$, and it is said to be *totally exceptional* if both conditions hold simultaneously. The relevance of the exceptional $\boldsymbol{\lambda}_\ell$ is that the analogue of the Fourier transform has a quite different behavior at them. To emphasize this point we write

$$\lambda_{\ell_j} = \frac{1}{4} - t_{\ell_j}^2 = \left(\frac{1}{2} + it_{\ell_j}\right) \left(\frac{1}{2} - it_{\ell_j}\right) \quad \text{with } t_{\ell_j} \in [0, \infty) \cup i(0, 1/2].$$

(Note that we slightly divert from [BGM11]). In this way, $t_{01} = t_{02} = i/2$ and $\boldsymbol{\lambda}_\ell$ is exceptional if $\Im t_{\ell_1}$ or $\Im t_{\ell_2}$ belong to $(0, 1/2)$. Although it is conjectured that there are no exceptional $\boldsymbol{\lambda}_\ell$ in the cases of arithmetic interest (this is the generalization of a famous conjecture due to Selberg [Sel65]), in principle there might be infinitely many such $\boldsymbol{\lambda}_\ell$. On the other hand, only finitely many can be totally exceptional (because the set $\{\boldsymbol{\lambda}_\ell\}$ is a discrete set) and the result of Kim and Shahidi [KS02] implies $\Im t_{\ell_1}, \Im t_{\ell_2} < 1/9$ for the lattices Γ in this paper.

As in the one-dimensional case, it turns out that $\{\psi_\ell\}_\ell$ does not span $L^2(\Gamma \backslash \mathbb{H}^2)$ if $\Gamma \backslash \mathbb{H}^2$ is not compact and a continuous spectrum corresponding to Eisenstein series enters into the game. The corresponding spectral theorem is technical in nature and we only need a particular case,

so we have limited its application to the proof of a single lemma. The reader preferring not to enter into the details of the proof can use Lemma 4.1 as a black box embodying the spectral theorem. We refer the reader to [BGM11] and [BM09, Ch.1] for more extensive comments on the spectral theorem (see also [HC68] for a more comprehensive theory).

Rather than the expansion of functions in $L^2(\Gamma \backslash \mathbb{H}^2)$, we need to expand a type of automorphic kernels. To introduce them it is convenient to extend the definition of u in (2.1) to \mathbb{H}^2 in the natural manner:

$$u(\mathbf{z}, \mathbf{w}) = (u(z_1, w_1), u(z_2, w_2)) \quad \text{for } \mathbf{z} = (z_1, z_2), \mathbf{w} = (w_1, w_2) \in \mathbb{H}^2.$$

Given an irreducible lattice Γ and $k : [0, \infty)^2 \rightarrow \mathbb{C}$ decaying rapidly enough at zero and infinity, for instance k compactly supported, we can construct an *automorphic kernel*

$$(2.4) \quad K(\mathbf{z}, \mathbf{w}) = \sum_{\gamma \in \Gamma} k(u(\gamma(\mathbf{z}), \mathbf{w})).$$

Using the fact that u is invariant by isometries, we deduce that K is actually automorphic in both variables, that is

$$K(\mathbf{z}, \mathbf{w}) = K(\gamma(\mathbf{z}), \mathbf{w}) = K(\mathbf{z}, \gamma(\mathbf{w})) \quad \text{for every } \gamma \in \Gamma.$$

3. AN ARITHMETIC LEMMA

To study the sum of $r(\lambda)r(\lambda+1)$ by analytic methods it is convenient to consider more general weighted sums

$$(3.1) \quad \mathcal{N}_k = \sum_{\lambda \in \mathcal{O}} r(\lambda)r(\lambda+1)k(\lambda, \lambda^\sigma) \quad \text{with } k : \mathbb{R}^2 \rightarrow \mathbb{C}.$$

A preliminary consideration is whether this sum actually makes sense when k decays rapidly enough. We state a general elementary result of this kind although, for the aims of this paper, we could restrict ourselves to compactly supported functions.

Lemma 3.1. *If $k(x, y) \ll (x + y + 1)^{-\alpha}$ with $\alpha > 3$ for $x, y \geq 0$ then \mathcal{N}_k is well-defined.*

Proof. Note first that if λ is a sum of two squares then so is λ^σ and if both are positive then we can restrict the sum to $\lambda, \lambda^\sigma > 0$.

If $\lambda = n + m\sqrt{D}$, expanding $\lambda = (a + b\sqrt{D})^2 + (c + d\sqrt{D})^2$ we see that $r(\lambda)$ counts the number of integral or half-integral solutions of

$$\begin{cases} a^2 + c^2 + D(b^2 + d^2) = n, \\ 2ab + 2cd = m. \end{cases}$$

The positivity of the first equation shows at once that $r(\lambda)$ is well-defined i.e., $r(\lambda) < \infty$. In fact using that the number of representations of an integer as a sum of two squares in \mathbb{Z} tends to zero when divided by any positive power [HW79, Th. 338], we have the trivial bound $r(\lambda) = O(n^{1+\epsilon})$ for any $\epsilon > 0$. Note also that, necessarily, in order to have a solution one must have $|m| < 4n$.

The inequality $2r(\lambda)r(\lambda + 1) \leq r^2(\lambda) + r^2(\lambda + 1)$ and the equation $\lambda + \lambda^\sigma = 2n$ reduce the assertion to proving that

$$\sum_{0 \leq m < 4n} \sum \frac{r^2(n + m\sqrt{D})}{n^\alpha} < \infty \quad \text{for } \alpha > 3.$$

Using the trivial bound we have

$$\sum_{0 \leq m < 4n} \sum \frac{r^2(n + m\sqrt{D})}{n^\alpha} \ll \sum_{0 \leq m < 4n} \sum \frac{r(n + m\sqrt{D})}{n^{\alpha-1-\epsilon}} \ll \sum'_{a,b,c,d} (a^2 + Db^2 + c^2 + Dd^2)^{1+\epsilon-\alpha}$$

where we disregard the value $a = b = c = d = 0$ in the last sum. It is plain that the latter series converges when $\alpha > 3 + \epsilon$, for instance by comparing with the integral $\int_{B'} \|\vec{x}\|^{-2-\delta} d\vec{x}$, $\delta > 0$, where B' is the exterior of the unit ball in \mathbb{R}^4 . \square

The key point to apply spectral methods is to translate \mathcal{N}_k into an automorphic kernel (2.4). The argument is an adaptation of that in the 1-dimensional case in [Iwa02, Cor.12.2].

Lemma 3.2. *Consider the lattice in $\mathrm{PSL}_2(\mathbb{R})^2$ defined as*

$$\Gamma = \{(\gamma, \gamma^\sigma) : \gamma \in M/\{\pm \mathrm{Id}\}\} \quad \text{with} \quad M = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}) : a + d, b + c \in 2\mathcal{O} \right\}.$$

Then for k as in Lemma 3.1 we have

$$\mathcal{N}_k = 2 \sum_{\gamma \in \Gamma} k(u(\gamma(\mathbf{i}), \mathbf{i})) \quad \text{where } \mathbf{i} = (i, i).$$

Proof. The map

$$\begin{aligned} \mathcal{C} := \{(A, B, C, D) \in \mathcal{O}^4 : A^2 + B^2 = C^2 + D^2 + 1\} &\longrightarrow M \\ (A, B, C, D) &\longmapsto \begin{pmatrix} A + C & B + D \\ D - B & A - C \end{pmatrix} \end{aligned}$$

clearly establishes a bijection between \mathcal{C} and M . On the other hand, if τ denotes the last matrix a calculation proves $u(\tau(i), i) = C^2 + D^2$. Hence

$$\mathcal{N}_k = \sum_{(A,B,C,D) \in \mathcal{C}} k(C^2 + D^2, (C^2 + D^2)^\sigma) = \sum_{\tau \in M} k(u(\tau(i), i), u(\tau^\sigma(i), i)).$$

This proves the result because $\pm\tau$ give rise to the same element in Γ . Note that the sign changes do not affect the values of $\lambda = u(\tau(i), i)$ and $\lambda^\sigma = u(\tau^\sigma(i), i)$. \square

4. A ROUGH SPECTRAL BOUND

Here we state the consequence of the application of the spectral theorem to automorphic kernels in the form needed for our purposes. It will be convenient to classify the labels of the exceptional λ_ℓ in three sets:

$$\Lambda_0 = \{\ell : \lambda_\ell \text{ totally exceptional}\}, \quad \Lambda_j = \{\ell \notin \Lambda_0 : \Im t_{\ell_j} \in (0, 1/2)\} \quad j = 1, 2.$$

Lemma 4.1. *Let k_1 and k_2 be continuous functions $k_j : [0, \infty) \rightarrow \mathbb{C}$ with Selberg transforms h_j satisfying (2.2). Consider the automorphic kernel (2.4) with $k(x, y) = k_1(x)k_2(y)$. Define*

$$H_j = \sum_{n=0}^{\infty} 2^{2n} \sup_{t \in I_n} |h_j(t)| \quad \text{with } I_0 = [0, 2) \text{ and } I_n = [2^n, 2^{n+1}) \text{ for } n \geq 1$$

and

$$(4.1) \quad \mathcal{M} = \frac{h_1(i/2)h_2(i/2)}{|\Gamma \backslash \mathbb{H}^2|} + \sum_{\ell \in \Lambda_0} h_1(t_{\ell_1})h_2(t_{\ell_2})\bar{\psi}_{\ell}(\mathbf{z})\psi_{\ell}(\mathbf{w}).$$

Then we have

$$K(\mathbf{z}, \mathbf{w}) = \mathcal{M} + O_{\mathbf{z}, \mathbf{w}, \Gamma} \left(H_1 H_2 + H_1 \sup_{\ell \in \Lambda_2} |h_2(t_{\ell_2})| + H_2 \sup_{\ell \in \Lambda_1} |h_1(t_{\ell_1})| \right).$$

REMARK. The dependence of the O constant on \mathbf{z} and \mathbf{w} could be made explicit but it is irrelevant in our application. Of course if $\Lambda_0 = \emptyset$ the sum over $\ell \in \Lambda_0$ must be omitted in \mathcal{M} and if there are no exceptional eigenvalues, the same applies to the suprema over Λ_1 and Λ_2 in the error term.

Proof. The spectral expansion of K , the analogue of the Poisson summation formula, as given in [BGM11, (39)], reads

$$\begin{aligned} K(\mathbf{z}, \mathbf{w}) &= \sum_{\ell} h(\mathbf{t}_{\ell})\bar{\psi}_{\ell}(\mathbf{z})\psi_{\ell}(\mathbf{w}) + \\ &\quad + 2 \sum_{\kappa} c_{\kappa} \sum_{\mu \in \mathcal{L}_{\kappa}} \int_{(\mathbb{R}^+)^2} h(\mathbf{t} + \mu) \bar{E}(\kappa; i\mathbf{t}, i\mu; \mathbf{z}) E(\kappa; i\mathbf{t}, i\mu; \mathbf{w}) dt_1 dt_2 \end{aligned}$$

where $h(t_1, t_2) = h_1(t_1)h_2(t_2)$, κ runs over the finitely many inequivalent cusps, c_{κ} are positive constants, \mathcal{L}_{κ} is a lattice in \mathbb{R}^2 and E denotes the Eisenstein series. In the first sum the terms with $\ell \in \Lambda_0 \cup \{0\}$ contribute exactly as \mathcal{M} . Let $K^* = K - \mathcal{M}$, we have to prove that it is bounded by the error term in the statement. Using that $|ab| \leq (|a|^2 + |b|^2)/2$ we have for $\mathbf{v} = \mathbf{z}$ or $\mathbf{v} = \mathbf{w}$

$$(4.2) \quad |K^*(\mathbf{z}, \mathbf{w})| \leq \sum_{\ell \notin \Lambda_0 \cup \{0\}} |h(\mathbf{t}_{\ell})| |\psi_{\ell}(\mathbf{v})|^2 + \\ + 2 \sum_{\kappa} c_{\kappa} \sum_{\mu \in \mathcal{L}_{\kappa}} \int_{(\mathbb{R}^+)^2} |h(\mathbf{t} + \mu)| |E(\kappa; i\mathbf{t}, i\mu; \mathbf{v})|^2 dt_1 dt_2.$$

Now we need a form of Bessel's inequality that allows to bound, for a fixed $\mathbf{z} \in \mathbb{H}^2$ and every $(n_1, n_2) \in \mathbb{Z}_{\geq 0}^2$, the expression

$$S(n_1, n_2, \mathbf{z}) = \sum_{\ell \in \mathcal{X}_d} |\psi_{\ell}(\mathbf{z})|^2 + 2 \sum_{\kappa} c_{\kappa} \sum_{\mu \in \mathcal{L}_{\kappa}} \int_{\mathcal{X}_c} |E(\kappa; i\mathbf{t}, i\mu; \mathbf{z})|^2 dt_1 dt_2$$

where

$$\mathcal{X}_d = \{ \ell : t_{\ell_j} \in I_{n_j} \cup i(0, 1/2] \} \quad \text{and} \quad \mathcal{X}_c = \{ \mathbf{t} \in (\mathbb{R}^+)^2 : \pm(t_j + \mu_j) \in I_{n_j} \}.$$

The instance of Bessel's inequality we need is [BGM11, Th. 4.2]

$$(4.3) \quad S(n_1, n_2, \mathbf{z}) = O(2^{2n_1+2n_2}).$$

The intuitive interpretation is that ψ_ℓ and E behave as constants on average.

If we divide the integral in (4.2) into the dyadic pieces $\pm(t_j + \mu_j) \in I_{n_j}$ indicated by \mathcal{X}_c and, using the positivity, we apply (4.3) to each of them, we have that the last term in (4.2) contributes at most

$$2 \sum_{n_1, n_2=0}^{\infty} \sup_{t \in I_{n_1}} |h_1(t)| \sup_{t \in I_{n_2}} |h_2(t)| \sum_{\kappa} c_{\kappa} \sum_{\mu \in \mathcal{L}_{\kappa}} \int_{\mathcal{X}_c} |E(\kappa; i\mathbf{t}, i\boldsymbol{\mu}; \mathbf{v})|^2 dt_1 dt_2 = O(H_1 H_2).$$

The same argument works to get this bound for the contribution of the first term in the right-hand side of (4.2) when t_{ℓ_1} and t_{ℓ_2} are real. The remaining terms have $\ell \in \Lambda_1 \cup \Lambda_2$ and we can proceed in the same way keeping the supremum of h_j if $\ell \in \Lambda_j$. For instance, the terms with $\ell \in \Lambda_1$ contribute at most

$$\sup_{t_{\ell_1}} |h_1(t_{\ell_1})| \sum_{\ell \in \Lambda_1} \sum_{n_2=0}^{\infty} \sup_{t \in I_{n_2}} |h_2(t)| |\psi_{\ell}(\mathbf{v})|^2 \leq \sup_{t_{\ell_1}} |h_1(t_{\ell_1})| \sum_{n_2=0}^{\infty} \sup_{t \in I_{n_2}} |h_2(t)| S(0, n_2, \mathbf{v})$$

and the sum is $O(H_2)$ by (4.3).

Therefore, we have proved that $|K^*(\mathbf{z}, \mathbf{w})|$ is bounded by the error term appearing in the statement. \square

5. VOLUME COMPUTATIONS

The main term (4.1) in the spectral expansion depends on the volume of the fundamental region and it becomes closely related to the constant C_D appearing in the asymptotic formulas in Theorem 1.1 and Theorem 1.2.

Our aim in this section is to prove the following result:

Proposition 5.1. *Let Γ be the subgroup of the full Hilbert modular group $\Gamma_{\mathcal{O}}$ introduced in Lemma 3.2. The volume of a fundamental region of $\Gamma \backslash \mathbb{H}^2$ is given by*

$$|\Gamma \backslash \mathbb{H}^2| = (2 - \chi(2) + 2\chi(4)) \frac{\pi^2}{\Delta} \sum_{n=1}^{\Delta} n^2 \chi(n)$$

where χ is the character corresponding to the Kronecker symbol $\left(\frac{\Delta}{\cdot}\right)$ and Δ is the discriminant of $\mathbb{Q}(\sqrt{D})$ i.e., $\Delta = D$ if $4 \nmid D - 1$ and $\Delta = 4D$ otherwise.

To prove it we are going to use an old result due to Siegel [Sie36b] and the computation of the index $[\Gamma_{\mathcal{O}} : \Gamma]$. We give indeed an explicit description of the representatives of the subgroups.

The elements of the full Hilbert modular group and its subgroups are pairs of matrices related by the real conjugation. This redundant presentation is important when the action on \mathbb{H}^2 is considered but from the group theoretical point of view we get an isomorphic group dropping the second matrix in the pair. For the sake of brevity, in the next result and in the rest of the section we identify (γ, γ^σ) and γ when the action on \mathbb{H}^2 is irrelevant.

Proposition 5.2. *Define*

$$T_u = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \omega = \frac{1 + \sqrt{D}}{2}, \quad \bar{\omega} = \frac{1 - \sqrt{D}}{2}, \quad \eta = \begin{cases} 1 + \sqrt{D} & \text{if } 4 \mid D - 3, \\ \sqrt{D} & \text{if } 4 \mid D - 2 \end{cases}$$

and

$$C_1 = \{\text{Id}\} \cup \{T_u\}_{u \in \Omega}, \quad C_2 = \{ST_u\}_{u \in \Omega}, \quad C_3 = \{T_u ST_v\}_{v \in \Omega^*}^{u \in \Omega}, \quad C_4 = \{ST_v ST_v\}_{v \in \Omega^*}$$

where $\Omega = \{1, \omega, \bar{\omega}\}$ and $\Omega^* = \{\omega, \bar{\omega}\}$. Then a complete set of representatives of the cosets $\Gamma_{\mathcal{O}}/\Gamma$ is given by

$$\mathcal{R}_D = \begin{cases} \{\text{Id}, T_1, T_\eta, T_{\eta+1}, ST_1, ST_{\eta+1}\} & \text{if } 4 \nmid D - 1, \\ C_1 \cup C_2 \cup \{T_1 ST_\omega, T_1 ST_{\bar{\omega}}\} & \text{if } 8 \mid D - 1, \\ C_1 \cup C_2 \cup C_3 \cup C_4 & \text{if } 8 \mid D - 5. \end{cases}$$

Hence the index in these three cases is respectively, 6, 9 and 15. This can be written in an artificial but compact way with the Kronecker symbol.

Corollary 5.3. *We have $[\Gamma_{\mathcal{O}} : \Gamma] = 6 - 3\chi(2) + 6\chi(4)$ with χ as in Proposition 5.1.*

We divide the proof of Proposition 5.2 according whether 8 divides $D-5$ or not. In this second case we benefit from a simple description of the group Γ given in the following result which allows a substantial reduction in the computations. In its proof and in that of Proposition 5.2 we will use that for $4 \nmid D-1$, we have $\eta^2 \in 2\mathcal{O}$ and when writing each $x \in \mathcal{O}$ as $x = x_1 + x_2\eta$, $x_1, x_2 \in \mathbb{Z}$ we have $x \in 2\mathcal{O}$ if and only if x_1 and x_2 are even.

Lemma 5.4. *If $8 \nmid D-5$ we have $\Gamma = C^{-1}\Gamma_0(2\mathcal{O})C$ where $C = ST_1$ and $\Gamma_0(2\mathcal{O})$ is the subgroup of matrices in $\Gamma_{\mathcal{O}}$ with lower left entry in $2\mathcal{O}$. Moreover, two matrices in $\Gamma_{\mathcal{O}}$ belong to the same left coset of $\Gamma_0(2\mathcal{O})$ if and only if the determinant of the matrix formed by their first columns belongs to $2\mathcal{O}$.*

Proof. Note the computations

$$C \begin{pmatrix} a & b \\ c & d \end{pmatrix} C^{-1} = \begin{pmatrix} * & * \\ a+c-b-d & * \end{pmatrix} \quad \text{and} \quad C^{-1} \begin{pmatrix} a & b \\ 2c & d \end{pmatrix} C = \begin{pmatrix} b+d & -a+b-2c+d \\ -b & a-b \end{pmatrix}.$$

The first one shows $\Gamma \subset C^{-1}\Gamma_0(2\mathcal{O})C$ because if $a+d, b+c \in 2\mathcal{O}$ then $a+c-b-d \in 2\mathcal{O}$.

The second computation shows that $\Gamma \supset C^{-1}\Gamma_0(2\mathcal{O})C$, it reduces to check that $ad-1 \in 2\mathcal{O}$ implies $a+d \in 2\mathcal{O}$. We consider two cases depending whether or not 8 divides $D-1$.

If $8 \mid D-1$, write $a = a_1 + a_2\omega$ and $d = d_1 + d_2\omega$ and note $\omega^2 - \omega \in 2\mathcal{O}$. Expanding ad we get that $ad-1 \in 2\mathcal{O}$ if and only if $2 \nmid a_1d_1$ and $2 \mid a_1d_2 + a_2d_1 + a_2d_2$ or equivalently if a_1, d_1 are both odd and a_2, d_2 are both even. Hence $a+d \in 2\mathcal{O}$.

If $8 \nmid D-1$, write $a = a_1 + a_2\eta$ and $d = d_1 + d_2\eta$. Expanding ad we see that $ad-1 \in 2\mathcal{O}$ implies $2 \nmid a_1d_1 + a_2d_2$ and $2 \mid a_1d_2 + a_2d_1$ in particular a_1 and a_2 have different parity. In fact we can assume $2 \nmid a_1$ (by the symmetry $a_1 \leftrightarrow a_2, d_1 \leftrightarrow d_2$) then $2 \mid a_2$ and we conclude $2 \nmid d_1, 2 \mid d_2$ that gives $a+d \in 2\mathcal{O}$ as expected.

Finally, $\gamma_1, \gamma_2 \in \Gamma_{\mathcal{O}}$ belong to the same coset if and only if $\gamma_2^{-1}\gamma_1 \in \Gamma_0(2\mathcal{O})$ and the last part of assertion in the result reduces to write the formula for the lower left entry of this product. \square

of Proposition 5.2 for $8 \nmid D - 5$. We check first that different elements in \mathcal{R}_D represent different cosets. A calculation shows that $T_u^{-1}ST_v \in \Gamma$ for $u, v \in \mathcal{O}$ if and only if $u - v, v^2 \in 2\mathcal{O}$. Then T_u and ST_v are in different cosets when $v \in \{1, \eta + 1\}$ if $4 \nmid D - 1$ and when $v \in \{1, \omega, \bar{\omega}\}$ if $8 \mid D - 1$. Clearly T_u and T_v belong to different cosets when $u - v \notin 2\mathcal{O}$, since $T_u^{-1}T_v = T_{v-u}$, and the same applies to ST_u and ST_v . It only remains to check that in the case $8 \mid D - 1$ the elements T_1ST_{ω} and $T_1ST_{\bar{\omega}}$ do not share coset with the other elements. As T_u, ST_{ω} and $ST_{\bar{\omega}}$ are in different cosets, the same holds for T_u and T_1ST_v with $v \in \{\omega, \bar{\omega}\}$. Writing $u = a + bv$ and using $v^2 - v \in 2\mathcal{O}$, after a calculation $(ST_u)^{-1}T_1ST_v \in \Gamma$ imposes $2 \mid a + 1, 2 \mid b + 1$ and $2 \mid a + b + 1$ which leads to a contradiction.

We focus firstly on the case $4 \nmid D - 1$. By Lemma 5.4 we have to prove that for each element in $\Gamma_{\mathcal{O}}$ there is an element in $C\mathcal{R}_DC^{-1}$ belonging to the same coset of $\Gamma_0(2\mathcal{O})$. The first column of the matrices in $C\mathcal{R}_DC^{-1}$ is given by the following vectors, except for adding to the coordinates elements of $2\mathcal{O}$,

$$\text{Id} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T_1 \rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad T_{\eta} \rightarrow \begin{pmatrix} 1 \\ \eta \end{pmatrix}, \quad T_{\eta+1} \rightarrow \begin{pmatrix} 1 \\ \eta + 1 \end{pmatrix}, \quad ST_1 \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad ST_{\eta+1} \rightarrow \begin{pmatrix} \eta \\ \eta + 1 \end{pmatrix}.$$

With the criterion given at the end of Lemma 5.4 it is enough to prove that if we add to these vectors a column with elements $a, b \in \mathcal{O}$ at least one of the corresponding determinants is in $2\mathcal{O}$. Clearly we can assume $a, b \in \{0, 1, \eta, \eta + 1\}$. Note that the vectors corresponding to Id and ST_1 take care of all the cases with a or b zero. Then there are nine cases to be considered. Three of them have $a = b$ and the determinant with the second vector is zero. By the same reason, we can also disregard the three cases in which a and b form the vectors corresponding to $T_{\eta}, T_{\eta+1}$ and $ST_{\eta+1}$. The remaining cases are $(a, b) = (\eta + 1, \eta), (\eta + 1, 1)$ and $(\eta, 1)$ and we get determinants in $2\mathcal{O}$ using the vectors corresponding respectively to $T_{\eta}, T_{\eta+1}$ and $ST_{\eta+1}$.

We deal now with the case $8 \mid D - 1$. To prove that for each element in $\Gamma_{\mathcal{O}}$ there exists an element in $C\mathcal{R}_DC^{-1}$ in the same coset of $\Gamma_0(2\mathcal{O})$ we proceed as before. This case is much simpler and no calculations are needed because after excluding the cases $b = 0, a = 0$ and $a = b$ using as above respectively Id, ST_1 and T_1 , we have only six possibilities with $a, b \in \{1, \omega, \bar{\omega}\}$ and they are all covered since each remaining element in $C\mathcal{R}_DC^{-1}$ treats at least the case corresponding to its own first column. \square

If $8 \mid D - 5$ the group Γ is not a conjugate of $\Gamma_0(2\mathcal{O})$. In this case, and actually also when $8 \mid D - 1$, there is a simple set of generators of $\Gamma_{\mathcal{O}}$. Considering some relations among them it is possible to simplify any word to one of the representatives indicated below multiplied by an element of Γ . The drawback of this method is that it leads to distinguish a number of cases that require somewhat tedious calculations. The advantage is that it gives a unified treatment of the case $4 \mid D - 1$ (see the remarks after the proof) and it potentially works for other subgroups.

of Proposition 5.2 for $8 \mid D - 5$. A result due to Vaseršteĭn [Vas72] (see also [Lie81], [May07, §1.2.2] and [Eve16, §5.1]) assures that the group $\Gamma_{\mathcal{O}}$ is generated by S , T_1 and T_{ω} . We note that $S^2 = -\text{Id}$ and $S \in \Gamma$. It is clear that, modulo Γ we can write any element $g \in \Gamma_{\mathcal{O}}$ as an alternating product of factors equal to S and T_u with $u \in \mathcal{O}$, that we will call generically a word. We employ the usual notation $g_1 \sim g_2$, or g_1 and g_2 equivalent modulo Γ , to mean that $g_1, g_2 \in \Gamma_{\mathcal{O}}$ belong to the same left coset i.e., $g_2^{-1}g_1 \in \Gamma$. Note that $gS \sim g$ and then we can always consider words with $T_u \notin \Gamma$ to the right.

By the multilinear properties of the multiplication of matrices, replacing in a product T_u by T_{u+w} with w in $2\mathcal{O}$ changes the entries of the final result in elements in this ideal. Recalling the definition of Γ , it is easy to see that two matrices in $\Gamma_{\mathcal{O}}$ with entries differing in elements of $2\mathcal{O}$ belong to the same coset. Hence it is enough to consider products involving translations T_u with $u \in \Omega$ because for any $u \in \mathcal{O}$ we can find $w \in 2\mathcal{O}$ such that $u - w \in \{0\} \cup \Omega$. Note that any word of length one is equivalent to exactly a one element of C_1 .

A calculation shows

$$(5.1) \quad T_v^{-1}ST_uST_v = \begin{pmatrix} -1 - uv & -uv^2 \\ u & uv - 1 \end{pmatrix}.$$

For $v = 1$ it lies in Γ . This implies that for any $g \in \Gamma_{\mathcal{O}}$,

$$(5.2) \quad gT_uST_1 = gST_1(T_{-1}ST_uST_1) \sim gST_1.$$

Hence any element of the form $gT_1 \notin \Gamma$ is equivalent to T_1 or ST_1 .

The set C_2 contains representatives of all the words of length 2 and they are clear nonequivalent. We have

$$(5.3) \quad T_uST_v = \begin{pmatrix} -u & 1 - uv \\ -1 & -v \end{pmatrix}$$

and it is easy to deduce that $C_1 \cup C_2$ does not contain equivalent elements. As shown before, the words T_uST_v can be simplified if $v = 1$, hence they are all equivalent to some element in $C_1 \cup C_2 \cup C_3$. Noticing

$$(5.4) \quad \omega\bar{\omega} - 1 \in 2\mathcal{O} \quad \text{and} \quad \omega^2 - \bar{\omega} \in 2\mathcal{O},$$

it follows from (5.3) that $C_1 \cup C_3$ does not contain equivalent elements and by (5.1) (recall $S^{-1} = -S$) any element in C_2 is not equivalent to an element in C_3 .

For the words of length four ST_uST_v we could consider in principle $u \in \Omega$, $v \in \Omega^*$ which makes six possibilities but some calculations show that $(T_{\bar{\omega}}ST_{\omega})^{-1}ST_{\bar{\omega}}ST_{\omega}$ and $(T_1ST_{\bar{\omega}})^{-1}ST_1ST_{\omega}$ are respectively

$$\begin{pmatrix} \omega + \bar{\omega}(\omega\bar{\omega} - 1) & \omega^2 + (\omega\bar{\omega} - 1)^2 \\ -1 - \bar{\omega}^2 & -\omega - \bar{\omega}(\omega\bar{\omega} - 1) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2\bar{\omega} - 1 & 2\omega\bar{\omega} - \omega - \bar{\omega} + 1 \\ -2 & 1 - 2\omega \end{pmatrix},$$

that belong to Γ using (5.4). Taking conjugates we deduce

$$(5.5) \quad ST_{\bar{\omega}}ST_{\omega} \sim T_{\bar{\omega}}ST_{\omega}, \quad ST_{\omega}ST_{\bar{\omega}} \sim T_{\omega}ST_{\bar{\omega}}, \quad ST_1ST_{\omega} \sim T_1ST_{\bar{\omega}}, \quad ST_1ST_{\bar{\omega}} \sim T_1ST_{\omega}.$$

Then the only possibilities for elements not equivalent to those in $C_1 \cup C_2 \cup C_3$ are the ones considered in C_4 .

Assuming that the elements in \mathcal{R}_D are not equivalent, we are going to prove that they form a complete set of representatives. This follows by an inductive argument if we prove that the words with length 5 are equivalent to shorter words because we have seen that the words of length at most 4 are equivalent to elements in \mathcal{R}_D . By (5.5) and (5.2) it is enough to consider words with $ST_\omega ST_\omega$ or $ST_{\bar{\omega}} ST_{\bar{\omega}}$ to the right. In fact, taking real conjugates if necessary, it is enough to prove that $T_u ST_\omega ST_\omega \sim ST_\omega ST_\omega$, or equivalently

$$(5.6) \quad (ST_\omega ST_\omega)^{-1} T_u ST_\omega ST_\omega = \begin{pmatrix} u(\omega^2 - 1)\omega + \omega - \omega^2 + 1 & u(\omega^2 - 1)^2 \\ -u\omega^2 & -u\omega(\omega^2 - 1) + 1 \end{pmatrix} \in \Gamma.$$

Using (5.4) we have that $\omega - \omega^2 + 1 \in 2\mathcal{O}$ and then the trace of this matrix belongs to $2\mathcal{O}$. In the same way, $(\omega^2 - 1)^2 - \omega^2 = \omega^4 - 3\omega^2 + 1$ differs from $\omega - 3\bar{\omega} + 1 \in 2\mathcal{O}$ in an element of $2\mathcal{O}$. Hence the matrix is in Γ .

It remains to prove that \mathcal{R}_D does not contain equivalent elements. Recall that we knew that this was the case for $C_1 \cup C_2 \cup C_3$. Note that $(ST_{\bar{\omega}} ST_{\bar{\omega}})^{-1} ST_\omega ST_\omega = T_{\bar{\omega}} ST_1 ST_\omega$ which is not in Γ by (5.1). Then $ST_\omega ST_\omega \not\sim ST_{\bar{\omega}} ST_{\bar{\omega}}$. By real conjugation we can restrict ourselves to check that $g = ST_\omega ST_\omega$ is not equivalent to any any element in $C_1 \cup C_2 \cup C_3$. From (5.6), $T_u g \sim g$ and then g is not equivalent to any element in $C_1 \cup C_3$. Clearly $h^{-1}g \in C_2 \cup C_3$ for $h \in C_2$, then $g \not\sim h$ also in this case and the proof is complete. \square

As mentioned before, it is possible to extend the previous proof to cover the case $8 \mid D - 1$. We now sketch the alternative proof following the same lines as in the case when $8 \mid D - 5$.

The calculation (5.1) does not depend on the properties of D and we can deduce as before that gT_1 is equivalent to Id , T_1 or ST_1 . In fact (5.1) implies

$$(5.7) \quad T_u ST_1 \sim ST_1, \quad T_\omega ST_\omega \sim ST_\omega \quad \text{and} \quad T_{\bar{\omega}} ST_{\bar{\omega}} \sim ST_{\bar{\omega}}.$$

The two latter establish the main difference with the case $8 \mid D - 5$ and are due to the fact that (5.4) must be replaced by $\omega\bar{\omega} \in 2\mathcal{O}$ and $\omega^2 - \bar{\omega} + 1 \in 2\mathcal{O}$ when $8 \mid D - 5$. Multiplying to the left by T_1 one deduces (recall $\omega + 1 - \bar{\omega} \in 2\mathcal{O}$)

$$(5.8) \quad T_{\bar{\omega}} ST_\omega \sim T_1 ST_\omega \quad \text{and} \quad T_\omega ST_{\bar{\omega}} \sim T_1 ST_{\bar{\omega}}.$$

Hence any word of length at most 3 is equivalent to one of the elements in $C_1 \cup C_2 \cup \{T_1 ST_\omega, T_1 ST_{\bar{\omega}}\}$. These elements are inequivalent (this can be obtained essentially following the scheme of the previous proof, where a bigger set is considered).

To conclude that it is a full set of representatives we have to prove that any word of length 4 can be reduced to one of its elements. By (5.7) and (5.8) it is enough to consider $ST_1 ST_\omega$ and $ST_1 ST_{\bar{\omega}}$. The following calculation shows $ST_1 ST_\omega \sim T_1 ST_{\bar{\omega}}$

$$(T_1 ST_{\bar{\omega}})^{-1} ST_1 ST_\omega = \begin{pmatrix} 2\bar{\omega} - 1 & 2\omega\bar{\omega} \\ -2 & 1 - 2\omega \end{pmatrix}$$

and by conjugation $ST_1 ST_{\bar{\omega}} \sim T_1 ST_\omega$.

Summarizing, the relations (5.5) were not enough to exclude words of length 4 when $8 \mid D - 5$ while the two last formulas in (5.7) cause the collapse of the potential new possibilities in the case $8 \mid D - 1$.

of *Proposition 5.1*. Siegel proved a closed formula [Sie36b, (19)] for $|\Gamma_{\mathcal{O}} \backslash \mathbb{H}^d|$ when \mathcal{O} is the ring of integers of a totally real number field. In the quadratic case, it reads

$$(5.9) \quad |\Gamma_{\mathcal{O}} \backslash \mathbb{H}^2| = \frac{2}{\pi^2} \Delta^{3/2} \zeta_D(2)$$

where ζ_D is the Dedekind zeta function of $\mathbb{Q}(\sqrt{D})$. We have $\zeta_D(2) = \zeta(2)L(2, \chi) = \frac{1}{6}\pi^2 L(2, \chi)$ with L the Dirichlet L -function associated to the character χ . By the functional equation of L in its asymmetric form [MV07, Cor.10.9] (note that χ is primitive and even)

$$\zeta_D(2) = -\frac{1}{3}\pi^4 \Delta^{-3/2} L(-1, \chi).$$

By Proposition 5.2, the volume of the fundamental region for Γ is $6 - 3\chi(2) + 6\chi(4)$ times that for $\Gamma_{\mathcal{O}}$, hence

$$|\Gamma \backslash \mathbb{H}^2| = -2\pi^2(2 - \chi(2) + 2\chi(4))L(-1, \chi).$$

Now we appeal to the known formula [Was97, Th.4.2] (as an aside, we note that the evaluation of $L(1 - n, \chi)$ plays an important role in the definition of p -adic L -functions)

$$L(-1, \chi) = -\frac{\Delta}{2} \sum_{n=1}^{\Delta} \chi(n) B_2(n/\Delta) \quad \text{with} \quad B_2(x) = x^2 - x + \frac{1}{6}.$$

As χ is even,

$$\sum_{n=1}^{\Delta} n\chi(n) = \sum_{n=1}^{\Delta} (\Delta - n)\chi(n)$$

and the sum of both sides equal $\Delta \sum_{n=1}^{\Delta} \chi(n) = 0$, hence both vanish. Then we can replace $B_2(n/\Delta)$ by $(n/\Delta)^2$. \square

6. PROOF OF THE MAIN RESULTS

After Lemma 3.2 and Lemma 4.1, the proof of Theorem 1.1 and Theorem 1.2 boils down to approximate sharp cuts by smooth kernels and estimating Selberg transforms. Following [Cha96], with an hyperbolic version of a classical Euclidean device, it is possible to reduce the whole problem to estimate the Selberg transform of the characteristic function of an interval (which is done in [Cha96, Lemma 2.4]). With this idea in mind we write χ_V and h_V respectively for the characteristic function of $[0, V]$ and its Selberg transform

$$\chi_V(x) = \begin{cases} 1 & \text{if } 0 \leq x < V, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad h_V(t) = \int_{\{u(z,i) \leq V\}} y^{1/2+it} d\mu(z).$$

The slow decay of h_V would cause serious convergence problems when applying the spectral theorem. The idea introduced in [Cha96] is to replace χ_V by a manageable approximation in such a way that its Selberg transform is like a product of two h_V 's at different scales. It doubles

the decay without requiring to estimate new transforms. The key argument is a simple one and it is summarized in the following result.

Lemma 6.1. *For $0 < v < V$ consider the function $F : \mathbb{H}^2 \rightarrow \mathbb{R}$ given by*

$$F(z, w) = \frac{1}{4\pi v} \int_{\mathbb{H}} \chi_V(u(z, \zeta)) \chi_v(u(\zeta, w)) d\mu(\zeta).$$

Then there exists $f : [0, \infty) \rightarrow \mathbb{R}$ such that $F(z, w) = f(u(z, w))$ and

$$\chi_{V^-} \leq f \leq \chi_{V^+} \quad \text{where} \quad V^\pm = \left(\sqrt{V(1+v)} \pm \sqrt{v(1+V)} \right)^2.$$

Moreover the Selberg transform of f is $h_V h_v$.

Proof. Note that $F(\gamma z, \gamma w) = F(z, w)$ for $\gamma \in \text{PSL}_2(\mathbb{R})$ because $d\mu$ is the invariant measure, then $F(z, w)$ only depends on $\rho(z, w)$ and it assures the existence of f . Using geodesic polar coordinates [Iwa02, (1.17)] it is plain that $\int_{\mathbb{H}} \chi_v(u(\zeta, w)) d\mu(\zeta) = 4\pi v$ (the area of a hyperbolic circle) and hence $0 \leq f \leq 1$.

Take $\tilde{V}, \tilde{v} > 0$ such that $2\tilde{V} + 1 = \cosh \tilde{V}$ and $2\tilde{v} + 1 = \cosh \tilde{v}$. Then by (2.1) we can write

$$f(u(z, i)) = \frac{1}{4\pi v^2} \int_{\mathbb{H}} \chi_{\tilde{V}}(\rho(z, \zeta)) \chi_{\tilde{v}}(\rho(\zeta, i)) d\mu(\zeta).$$

By the triangle inequality, if $\rho(z, i) \geq \tilde{V} + \tilde{v}$ the integral vanishes. As $0 \leq f \leq 1$, we can rewrite this as $f((\cosh x - 1)/2) \leq \chi_{\tilde{V} + \tilde{v}}(x)$ for $x \geq 0$ that means $f \leq \chi_{V^+}$ with $2V^+ + 1 = \cosh(\tilde{V} + \tilde{v})$ and the addition formula for \cosh gives the claimed formula for V^+ . In the same way, if $\rho(z, i) \leq \tilde{V} - \tilde{v}$ then $\rho(\zeta, i) < \tilde{v}$ implies $\rho(z, \zeta) < \tilde{V}$ and consequently we can omit $\chi_{\tilde{V}}$ to get $f(u(z, i)) = 1$. This can be rephrased as $f((\cosh x - 1)/2) \geq \chi_{\tilde{V} - \tilde{v}}(x)$ for $x \geq 0$ and proceeding as before, $f \geq \chi_{V^-}$.

The last part of the statement reduces to an application of Fubini's theorem using that

$$\int_{\mathbb{H}} \chi_V(u(z, w)) (\Im z)^{1/2+it} d\mu(\zeta) = h_V(t) (\Im w)^{1/2+it}.$$

See [CRRC13, Lemma 2.2] for the details and a general result. \square

Theorem 1.1 and Theorem 1.2 will be proved by an application of Lemma 4.1, for different choices of the parameters v_1 and v_2 in the following result.

Proposition 6.2. *Let $V_j > 0$ and $0 < v_j < C \min(1, V_j)$, $j = 1, 2$ where $0 < C < 1$ is an absolute constant. Then there exists $V'_j > 0$ with $V'_j = V_j + O(V_j v_j^{1/2} + (V_j v_j)^{1/2})$ such that*

$$N_D(V_1, V_2) = 2 \sum_{(\gamma, \gamma^\sigma) \in \Gamma} k_1(u(\gamma i, i)) k_2(u(\gamma^\sigma i, i))$$

where the Selberg transform of k_j is $(4\pi v_j)^{-1} h_{V'_j} h_{v_j}$.

REMARK. The constant C can be substituted by an explicit numerical value by following the steps in the proof, but this value is unimportant for our arguments, so we shall not get into this calculation.

Proof. Let us abbreviate $M_j = \max(1, V_j^{-1/2})$; then $0 < C^{-1/2}v_j^{1/2}M_j < 1$.

Take in Lemma 6.1 $v = v_j$ and $V = V_j(t)$ with $V_j(t) = V_j + tC^{-1/4}V_jM_jv_j^{1/2}$ and $t \in [-1, 1]$. For t in this interval, the hypothesis $0 < v < V$ is satisfied choosing C small enough because $C^{-1/4}M_jv_j^{1/2} \leq C^{1/4} \rightarrow 0$ when $C \rightarrow 0^+$. Let f_{t_j} be the corresponding function f in Lemma 6.1, which ensures that $f_{-1_j} \leq \chi_{V_j} \leq f_{1_j}$ under

$$(r_j^+ - s_j^+)^2 \geq V_j \quad \text{and} \quad (r_j^- + s_j^-)^2 \leq V_j \quad \text{with} \quad \begin{cases} r_j^\pm = \sqrt{V_j(\pm 1)(1 + v_j)}, \\ s_j^\pm = \sqrt{v(1 + V_j(\pm 1))}. \end{cases}$$

We have

$$V_j^{-1/2}r_j^\pm = 1 \pm \frac{1}{2}C^{-1/4}M_jv_j^{1/2} + O(C^{-1/2}M_j^2v_j + v_j) \quad \text{and} \quad V_j^{-1/2}s_j^\pm = O(M_jv_j^{1/2})$$

hence the inequalities hold choosing C small enough.

Consequently, defining K_t as the automorphic kernel (2.4) that corresponds to $k_t(x, y) = f_{t_1}(x)f_{t_2}(y)$, we have by Lemma 3.2

$$2K_{-1}(\mathbf{i}, \mathbf{i}) \leq N_D(V_1, V_2) \leq 2K_1(\mathbf{i}, \mathbf{i}).$$

The intermediate value theorem implies $N_D(V_1, V_2) = 2K_{t_0}(V_1, V_2)$ for some $t_0 \in [-1, 1]$ and the result follows with $V_j' = V_j(t_0)$. \square

For the sake of completeness, we include here the estimates we need for the Selberg transform. A more precise result with asymptotic formulas is given in [Cha96].

Lemma 6.3. *The Selberg transform h_V of the characteristic function of $[0, V]$ satisfies for $t \geq 0$*

$$(6.1) \quad h_V(t) \ll \begin{cases} V^{1/2}(1+t)^{-3/2} & \text{if } t \geq 1, V \geq 1 \\ V^{1/2} \log(2V) & \text{if } t \leq 1, V \geq 1 \end{cases} \quad \text{and} \quad h_V(t) \ll V(1+t^2V)^{-3/4} \quad \text{if } V \leq 1,$$

and for t pure imaginary $0 < \Im t \leq 1/2$ we have

$$(6.2) \quad h_V(t) \ll V^{1/2+|t|} \min(|t|^{-1}, \log(2V)) \quad \text{if } V \geq 1 \quad \text{and} \quad h_V(t) \ll V \quad \text{if } V \leq 1.$$

Proof. This is a consequence of Lemma 2.4 in [Cha96]. \square

The logarithm appearing in these estimates cannot be avoided when V is large and t close to 0. It is reflected on some average results for the hyperbolic circle problem [PR94].

Now, we are ready to prove our main results.

of Theorem 1.1. Take in Proposition 6.2 $v_1 = v_2 = (V_1V_2)^{-1/2}$, hence we have $V_1' = V_1 + O(V_1^{3/4}V_2^{-1/4})$ and $V_2' = V_2 + O(V_2^{3/4}V_1^{-1/4})$. With the notation of Lemma 4.1, for $t \in I_n$, (6.1) gives

$$(4\pi v_j)^{-1} |h_{V_j'}(t)h_{v_j}(t)| \ll \frac{V_j^{1/2}}{(1+t)^{3/2}(1+t^2(V_1V_2)^{-1/2})^{3/4}}$$

and a term $\log(V_j + 1)$ must be introduced for $t \in I_0$. Then $H_j \ll V_j^{1/2}(V_1 V_2)^{1/8}$. On the other hand, the $1/9$ bound for $\Im t_{\ell_j}$ from [KS02] implies by (6.2)

$$(4\pi v_j)^{-1} |h_{V'_j}(t_{\ell_j}) h_{v_j}(t_{\ell_j})| \ll V_j^{1/2+1/9} = V_j^{11/18}.$$

Consequently, the sum over $\ell \in \Lambda_0$ in Lemma 4.1 contributes $O((V_1 V_2)^{11/18})$ and, as $h_{V'_j}(i/2) = 4\pi V'_j$ and $h_{v_j}(i/2) = 4\pi v_j$, we obtain the main term (4.1) of the automorphic kernel

$$\mathcal{M} = \frac{16\pi^2 V'_1 V'_2}{|\Gamma \backslash \mathbb{H}^2|} + O((V_1 V_2)^{11/18}) = \frac{1}{2} C_D V_1 V_2 + O((V_1 V_2)^{3/4})$$

where we have used expression (1.5) and Proposition 5.1. On the other hand, the error term in Lemma 4.1 is

$$O(V_1^{1/2}(V_1 V_2)^{1/8} \cdot V_2^{1/2}(V_1 V_2)^{1/8} + V_1^{1/2}(V_1 V_2)^{1/8} \cdot V_2^{11/18} + V_2^{1/2}(V_1 V_2)^{1/8} \cdot V_1^{11/18})$$

and clearly this is $O((V_1 V_2)^{3/4})$. \square

of Theorem 1.2. Take in Proposition 6.2 $v_1 = (V_1 V_2^2)^{-1/2}$ and $v_2 = V_2^{-1/2}$, that implies $V'_1 = V_1 + O(V_1^{1/2} V_2^{-1})$ and $V'_2 = V_2 + O(V_1^{-1/4} V_2^{1/2})$. The Selberg transforms of $h_{V'_1}$, h_{v_1} and h_{v_2} admit bounds as in the previous proof that give

$$(4\pi v_1)^{-1} |h_{V'_1}(t) h_{v_1}(t)| \ll \frac{V_1^{1/2}}{(1+t)^{3/2}(1+t^2(V_1 V_2^2)^{-1/2})^{3/4}} \quad \text{for } t \in I_n, n \geq 1$$

and $H_1 \ll V_1^{1/2}(V_1 V_2^2)^{1/8} = V_1^{5/8} V_2^{1/4}$.

The difference is that for $h_{V'_2}$ we have to use the second bound in (6.1) because $V'_2 < 1$. Then in this case

$$(4\pi v_2)^{-1} |h_{V'_2}(t) h_{v_2}(t)| \ll \frac{V_2}{(1+t^2 V_2)^{3/4} (1+t^2 V_1^{-1/2})^{3/4}} \quad \text{for } t \in I_n, n \geq 0$$

which leads to $H_2 \ll V_1^{1/8} V_2^{1/4}$.

If there are no exceptional eigenvalues, the error term in Lemma 4.1 is simply $O(H_1 H_2) = O(V_1^{3/4} V_2^{1/2})$ and we get the first part of the result.

For the second part, note that we only need a bound for $\Im t_{\ell_j}$ when $j = 1$ because the second formula of (6.2) does not involve t . Under our assumption

$$(4\pi v_1)^{-1} |h_{V'_1}(t_{\ell_1}) h_{v_1}(t_{\ell_1})| \ll V_1^{1/2+c} \quad \text{and} \quad (4\pi v_2)^{-1} |h_{V'_2}(t_{\ell_2}) h_{v_2}(t_{\ell_2})| \ll V_2.$$

Then the error term in Lemma 4.1 includes two new terms following $V_1^{3/4} V_2^{1/2}$, namely, it is

$$O(V_1^{3/4} V_2^{1/2} + V_1^{5/8} V_2^{5/4} + V_1^{5/8+c} V_2^{1/4}) = O(V_1^{3/4} V_2^{1/2} + V_1^{5/8+c} V_2^{1/4}).$$

Finally, note that the finite sum in \mathcal{M} contributes $O(V_1^{1/2+c} V_2)$ that is absorbed by the error term. \square

7. APPENDIX. NUMERICAL CONSIDERATIONS

A problem to carry out numerical calculations related to (3.1) is that r is essentially a divisor function in $\mathcal{O}[i]$ and it seems hard to implement in an efficient way. Here we make some comments for the reader interested in doing direct computations and we display some of the data obtained in this way.

A simple calculation, already displayed in the proof of Lemma 3.1 shows that for $\lambda = x + y\sqrt{D} \in \mathcal{O}$ the value of $r(\lambda)$ is the number of solutions a, b, c, d of

$$(7.1) \quad \begin{cases} x = a^2 + Db^2 + c^2 + Dd^2, \\ y = 2ab + 2cd. \end{cases}$$

Here, $a + b\sqrt{D}, c + d\sqrt{D} \in \mathcal{O}$. Equivalently, we consider $a, b, c, d \in \mathbb{Z}$ if $D \equiv 1 \pmod{4}$ and we consider $2a, 2b, 2c, 2d, a - b, c - d \in \mathbb{Z}$ if $D \not\equiv 1 \pmod{4}$.

The solutions of (7.1) enjoy some symmetries that are useful for numerical calculations.

Proposition 7.1. *Given $\lambda = x + y\sqrt{D} \in \mathcal{O}$, let M_j and M_k^* be the number of solutions $a + b\sqrt{D}, c + d\sqrt{D} \in \mathcal{O}$ of (7.1) satisfying the mutually exclusive conditions indicated in this list:*

$$\begin{aligned} M_1 &\rightarrow 0 < c < a; & M_2 &\rightarrow 0 = c < a, d > 0; & M_3 &\rightarrow 0 < c = a, d < b; \\ M_4 &\rightarrow 0 = c = a, 0 < d < b; & M_1^* &\rightarrow a = b = 0; & M_2^* &\rightarrow a = c, b = d. \end{aligned}$$

Then $r(\lambda) = 8(M_1 + M_2 + M_3 + M_4) + 2(M_1^* + M_2^*)$.

Proof. We have a group of transformations acting on the solutions, $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, generated by

$$(a, b, c, d) \mapsto (-a, -b, c, d), \quad (a, b, c, d) \mapsto (a, b, -c, -d), \quad (a, b, c, d) \mapsto (c, d, a, b).$$

This action is free (fixed point free) except for the set of solutions S^* satisfying one of the following four sets of conditions:

$$a = b = 0 \quad \text{or} \quad c = d = 0 \quad \text{or} \quad a = c, b = d \quad \text{or} \quad a = -c, b = -d.$$

The solutions satisfying the first or the second set of conditions amount $2M_1^*$ and the rest of the conditions give $2M_2^*$ solutions. In the complement of S^* the action of G gives equivalence classes with eight elements and we can always select exactly one of these eight elements satisfying the conditions indicated for M_1, M_2, M_3 and M_4 . This follows imposing an ordering between $0, a$ and c when possible i.e., when there are not coincidences among them. In this way M_1 counts the ordered case and M_2, M_3 and M_4 the cases with coincidences. \square

For the numerical calculation of $N_D(V_1, V_2)$ note that $0 \leq \lambda < V_1$ and $0 \leq \lambda^c < V_2$ is equivalent to

$$0 \leq 2x < V_1 + V_2, \quad \max(-x, x - V_2) < y\sqrt{D} < \min(V_1 - x, x).$$

The first one readily imposes upper bounds on $|a|$, $|b|$, $|c|$ and $|d|$. If we let a , b , c and d vary in these ranges and, for each combination, we add 1 to the position x , y of a matrix with x and y as in (7.1), then this matrix will store the required values of $r(\lambda)$ and the evaluation of $r(\lambda)r(\lambda + 1)$ in (3.1) corresponds to a sum of products of adjacent entries.

Now Proposition 7.1 allows to reduce in a factor of approximately 1/8 the ranges of a , b , c and d to be considered. In the implementation the loop corresponding to M_1 takes the bulk of the calculations because the rest of the loops involve a lesser number of free variables.

The actual issue for direct computations of this kind is the limit of the fast access allocatable memory to store the matrix. The size of the stored elements can be reduced to one fourth noting that $r(\lambda) = r(\lambda^\sigma)$ and that y is always even by (7.1). With these reductions, when $D \not\equiv 1 \pmod{4}$ in the more demanding case $V_1 \approx V_2 \approx V$ (less sparse matrix), the memory to be allocated for the values of x and y is of around

$$\sum_{x=0}^V \frac{\min(V-x, x)}{2\sqrt{D}} \sim \frac{V^2}{8\sqrt{D}}$$

integers. When V_1 and V_2 are very unbalanced, meaning V_1 large and $V_2 < 1$, then y is determined by x and one only needs room for the around $V_1/2$ integer values of x . If $D \equiv 1 \pmod{4}$ these figures should be multiplied by 2.

We mention here a couple of tables for $D = 2$ that we have obtained by implementing the previous idea in an average PC with a simple C program.

As in the case of the Gauss circle problem, the error term in the asymptotic approximation of $N_2(V, V)$ oscillates and to get a reliable idea on the limits for a general upper bound, rather than picking large special values of V , it is more informative to consider

$$F(x) = \sup_{V < x} |N_2(V, V) - 8V^2|.$$

Extensive computations prove

| x | 5000 | 10000 | 15000 | 20000 | 25000 | 30000 | 35000 | 40000 |
|--------|--------|--------|--------|--------|--------|---------|---------|---------|
| $F(x)$ | 124508 | 383780 | 421116 | 440700 | 882044 | 1082980 | 1369084 | 1807652 |

As an aside, a possibility that we have not explored is to carry out numerical calculations via Lemma 3.2 with a description of the group in terms of words and relations (cf. [PR94]). It looks promising when $4 \mid D - 1$ because $\Gamma_{\mathcal{O}}$ has simple generators and Proposition 5.2 gives a full description of the cosets but when $4 \nmid D - 1$ the natural generators are linked to the class group [Eve16, §5.2] and the calculations for large values of D could be more difficult. In both cases one should have some control on $u(\gamma(\mathbf{i}), \mathbf{i})$ in terms of the length of γ .

In the case of unbalanced arguments of N_2 , the numerical experiments suggest that the asymptotic formula

$$N_2(V_1, V_2) \sim 8V_1V_2$$

in Theorem 1.2 could hold in some ranges beyond the condition $V_1 V_2^2 \rightarrow \infty$, but even for reasonably large V_1 the values of $N_2(V_1, V_2)$ are too small to draw trustworthy conjectures. For instance, if we define

$$G(V) = \frac{1}{8} V^{-1/2} N_2(V, V^{-1/2})$$

then we have

| | | | | | |
|--------------------|----------|----------|----------|----------|----------|
| V | 10000 | 20000 | 30000 | 40000 | 50000 |
| $N_2(V, V^{-1/2})$ | 836 | 1220 | 1476 | 1540 | 1924 |
| $G(V)$ | 1.045000 | 1.078337 | 1.065211 | 0.962500 | 1.075548 |

It is tempting to claim that $G(V)$ goes to 1 when $V \rightarrow \infty$.

We finish with some considerations about the constant C_D giving the limit of $N_D(V, V)/V^2$. Here it is a table of some values

| | | | | | | | | | | |
|-------|---|---|---|-----|---|------|-------|---------|-----------|------------|
| D | 2 | 3 | 5 | 6 | 7 | 101 | 1001 | 10001 | 100001 | 1000001 |
| C_D | 8 | 4 | 8 | 4/3 | 1 | 8/95 | 2/753 | 1/11616 | 4/1462371 | 1/11832936 |

As we mentioned before, and the table confirms, C_D decays as $D^{-3/2}$. This decay can be proved theoretically with explicit constants. Namely, we have

Proposition 7.2. *With the notation as in (1.5), we have*

$$\frac{192}{5\Delta^{3/2}} < C_D < \frac{240}{\Delta^{3/2}}.$$

This is in accordance with the previous table. In which the minimal value of $\Delta^{3/2} C_D$ is around 84.12 (reached at $D = 1001$) and the maximal value is around 181.02 (reached at $D = 2$).

Proof. The formula (5.9) due to Siegel for the volume of the fundamental region of the full Hilbert modular group and Corollary 5.3 imply

$$|\Gamma \backslash \mathbb{H}^2| = (2 - \chi(2) + 2\chi(4)) \frac{6}{\pi^2} \Delta^{3/2} \zeta(2) L(2, \chi).$$

Clearly

$$\frac{\zeta(4)}{\zeta(2)} = \prod_p (1 + p^{-2})^{-1} < L(2, \chi) < \prod_p (1 - p^{-2})^{-1} = \zeta(2).$$

Using $\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$ and comparing with Proposition 5.1,

$$\frac{1}{15} \Delta^{5/2} < \sum_{n=1}^{\Delta} n^2 \chi(n) < \frac{1}{6} \Delta^{5/2}.$$

Substituting in the definition of C_D , we complete the proof. □

Sharper bounds can be proved separating congruence classes modulo 8. For instance, if $8 \mid D - 1$ the previous argument leads to

$$\Delta^{3/2} C_D = \frac{32\pi^2}{3L(2, \chi)} > \frac{32\pi^2}{3\zeta(2)} = 64$$

and in fact $\Delta^{3/2}C_D \rightarrow 64$ if we choose $D = 4\prod_{p \leq N} p + 1$ with $N \rightarrow \infty$ to force $\chi(p) = 1$ for small primes. For example, taking $N = 17$ we have $D = 38798761$ and $\Delta^{3/2}C_D \approx 64.84$.

REFERENCES

- [Ber54] S. Berg. Part 1 of Selberg's Gottingen lectures: a rough transcription. Available in the website of the Institute for Advanced Study <https://publications.ias.edu/selberg>, 1954.
- [BGM11] R. W. Bruggeman, F. Grunewald, and R. J. Miatello. New lattice point asymptotics for products of upper half-planes. *Int. Math. Res. Not. IMRN*, (7):1510–1559, 2011.
- [BM09] R. W. Bruggeman and R. J. Miatello. Sum formula for SL_2 over a totally real number field. *Mem. Amer. Math. Soc.*, 197(919):vi+81, 2009.
- [BW17] J. Bourgain and N. Watt. Mean square of zeta function, circle problem and divisor problem revisited. arXiv:1709.04340 [math.AP], 2017.
- [Cha96] F. Chamizo. Some applications of large sieve in Riemann surfaces. *Acta Arith.*, 77(4):315–337, 1996.
- [Cha99] F. Chamizo. Correlated sums of $r(n)$. *J. Math. Soc. Japan*, 51(1):237–252, 1999.
- [CRRC13] F. Chamizo, D. Raboso, and S. Ruiz-Cabello. Exotic approximate identities and Maass forms. *Acta Arith.*, 159(1):27–46, 2013.
- [Eve16] L.M. Everhart. On generators of Hilbert modular groups of totally real number fields. Master's thesis, UNC Greensboro, Greensboro, 2016. Available online https://libres.uncg.edu/ir/uncg/f/Everhart_uncg_0154M_11992.pdf.
- [Gau81] K. F. Gauss. De nexu inter multitudinem classium, in quas formae binariae secundi gradus distribuntur, earumque determinantem. In *Werke*, volume 2, pages 269–291. Georg Olms Verlag, Hildesheim, New York, 1981.
- [Haf85] J. L. Hafner. Applications of spectral theory to number theory. *Rocky Mountain J. Math.*, 15(2):389–397, 1985. Number theory (Winnipeg, Man., 1983).
- [HC68] Harish-Chandra. *Automorphic forms on semisimple Lie groups*. Notes by J. G. M. Mars. Lecture Notes in Mathematics, No. 62. Springer-Verlag, Berlin-New York, 1968.
- [Hux96] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.
- [HW79] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. The Clarendon Press, Oxford University Press, New York, fifth edition, 1979.
- [Iwa02] H. Iwaniec. *Spectral methods of automorphic forms*, volume 53 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, second edition, 2002.
- [KS02] H. H. Kim and F. Shahidi. Cuspidality of symmetric powers with applications. *Duke Math. J.*, 112(1):177–197, 2002.
- [Lan93] S. Lang. *Real and functional analysis*, volume 142 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 1993.
- [Lie81] B. Liehl. On the group SL_2 over orders of arithmetic type. *J. Reine Angew. Math.*, 323:153–171, 1981.
- [May07] S. Mayer. *Hilbert Modular Forms for the Fields $\mathbb{Q}(\sqrt{5})$, $\mathbb{Q}(\sqrt{13})$ and $\mathbb{Q}(\sqrt{17})$* . PhD thesis, RWTH Aachen University, Aachen, 2007. Available online <http://publications.rwth-aachen.de/record/62415>.
- [MV07] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.
- [PR94] R. Phillips and Z. Rudnick. The circle problem in the hyperbolic plane. *J. Funct. Anal.*, 121(1):78–116, 1994.

- [Rau88] U. Rausch. Eine Formel für die Koeffizientensumme von Potenzreihen mit Anwendung auf das Kreis- und Kugelproblem in total reellen algebraischen Zahlkörpern. *Acta Arith.*, 50(4):381–404, 1988.
- [Sch62] W. Schaal. Übertragung des Kreisproblems auf reell-quadratische Zahlkörper. *Math. Ann.*, 145:273–284, 1961/1962.
- [Sch65] W. Schaal. On the expression of a number as the sum of two squares in totally real algebraic number fields. *Proc. Amer. Math. Soc.*, 16:529–537, 1965.
- [Sel65] A. Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.
- [Sie36a] C. L. Siegel. Mittelwerte arithmetischer Funktionen in Zahlkörpern. *Trans. Amer. Math. Soc.*, 39(2):219–224, 1936.
- [Sie36b] C. L. Siegel. The volume of the fundamental domain for some infinite groups. *Trans. Amer. Math. Soc.*, 39(2):209–218, 1936.
- [Sie74] W. Sierpiński. *Oeuvres choisies*. PWN—Éditions Scientifiques de Pologne, Warsaw, 1974. Tome I: Bibliographie, théorie des nombres et analyse mathématique.
- [Vas72] L. N. Vaseršteĭn. The group SL_2 over Dedekind rings of arithmetic type. *Mat. Sb. (N.S.)*, 89(131):313–322, 351, 1972.
- [Was97] L. C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1997.