Theta-lifts of Maaß waveforms[1]

Jens Bolte[2]

Abteilung für Theoretische Physik
Universität Ulm, Albert-Einstein-Allee 11
D-89069 Ulm
Germany

Stefan Johansson[3]

Department of Mathematics
Chalmers University of Technology
and Göteborg University
S-41296 Göteborg
Sweden

[1]Supported by Deutscher Akademischer Austauschdienst and Svenska Institutet
[2]Electronic address: bol@physik.uni-ulm.de
[3]Electronic address: sj@math.chalmers.se

Abstract

Let $\mathcal{O}$ be an arbitrary order in an indefinite quaternion division algebra over $\mathbb{Q}$. If $\mathcal{O}^1$ is the group of elements in $\mathcal{O}$ with norm equal to 1, and $\mathcal{H}$ the complex upper half-plane, then $X_{\mathcal{O}} := \mathcal{O}^1 \backslash \mathcal{H}$ is a compact Riemann surface. Furthermore, let $\Gamma_0(d) \subseteq SL_2(\mathbb{Z})$ be the Hecke congruence group of level d. Then $X_d := \Gamma_0(d) \backslash \mathcal{H}$ is a non-compact Riemann surface with finite volume. Let Δ be the hyperbolic Laplace-operator on $\mathcal{H}$. In certain situations, it is known that it is possible to relate the spectral resolutions of automorphic Laplacians in the compact case to the non-compact case. In this paper, we give an explicit construction of such correspondence in the case of Maaß waveforms. The construction uses Siegel theta functions and generalises the one in [8]. Furthermore, we prove that the theta-lifts commute with Hecke operators. Finally, we investigate to what extent the lifted forms are newforms or not.

1 Introduction

The spectral theory of automorphic Laplacians allows for a variety of different approaches. From a geometrical point of view it is natural to consider closed compact surfaces X endowed with Riemannian metrics of constant negative curvature, and the spectral resolution of the Laplacian on $L^2(X)$. In an arithmetical approach, the most natural example to start with is the modular group $SL_2(\mathbb{Z})$ and its automorphic Laplacian. The corresponding modular surface is then non-compact and has two branch points. The spectral resolution of the Laplacian has an absolutely continuous part in addition to the discrete one. In certain situations, it is possible to relate the spectral resolutions of automorphic Laplacians in compact and non-compact situations, respectively. To obtain this the compact surface has to come from a cocompact arithmetic Fuchsian group. The spectral correspondence thus achieved also extends to preserve not only Laplace eigenvalues, but also the eigenvalues of Hecke operators. From a representation theoretic point of view, the most general case of such spectral correspondences is covered by the Jacquet-Langlands correspondence, see [10]. A proof of this correspondence exploiting adelic theta functions was subsequently given by Shimizu [21]. In the classical context of Maaß waveforms and trace formulae, it seems however desirable to formulate spectral correspondences in a classical language in order to make them more explicit.

In this paper, we work out the details of a classical construction of the spectral correspondence when the cocompact arithmetic Fuchsian group is given by a unit group of an order in an arbitrary indefinite quaternion algebra over the rationals. In the case of certain types of orders, the relevant constructions were previously given by Hejhal [8]. We show that this procedure can be extended to arbitrary orders, and we also improve the result concerning the level of the congruence group. It turns out that there is a natural correspondence between the discriminant of the order and the level of the corresponding congruence group. Moreover, based on an investigation of Hecke operators, we are able to demonstrate that in case of non-maximal orders a proportion of the theta-lifts cannot be newforms. This is complementary to a result of Ribet [18], who proved for holomorphic forms of weight one that in the case of maximal orders the analogous lifts yield an isomorphism onto the newforms for an appropriate Hecke congruence group. In the case of Maaß waveforms, we are able to give a partial answer to the question whether the lifts yield newforms in the case of maximal orders. Also in the holomorphic situation, lifts from forms of half-integer weight to forms of integer weight were introduced by Shimura [22], and Niwa realised these lifts using theta functions [17]. A lift in the reverse direction, i.e. from integer weight to half-integer weight was subsequently provided by Shintani [23].

To be specific, let $\mathcal{H} := \{z = x + iy : \ x \in \mathbb{R}, \ y > 0\}$ be the complex upper half-plane. The hyperbolic metric $ds^2 = y^{-2}(dx^2 + dy^2)$ on $\mathcal{H}$ is of constant Gaussian curvature $K = -1$. The orientation preserving isometries of the Riemannian manifold $(\mathcal{H}, ds^2)$ are given by the fractional linear transformations

$$z \mapsto \frac{\alpha z + \beta}{\gamma z + \delta} , \tag{1.1}$$

where $z \in \mathcal{H}$ and $\left(\begin{smallmatrix}\alpha & \beta \\ \gamma & \delta\end{smallmatrix}\right) \in SL_2(\mathbb{R})$. A cofinite Fuchsian group Γ is a discrete subgroup of $SL_2(\mathbb{R})$ such that the orbit space $X_\Gamma := \Gamma\backslash\mathcal{H}$ is of finite volume,

$$\mathrm{vol}(X_\Gamma) = \int_{\mathcal{F}_\Gamma} d\mu(z) < \infty . \tag{1.2}$$

Here $d\mu(z) := y^{-2}dx\,dy$ denotes the hyperbolic volume form derived from the metric ds^2, and $\mathcal{F}_\Gamma \subset \mathcal{H}$ is a suitable fundamental domain for Γ. The finiteness condition (1.2) is equivalent to Γ having a finite number κ of inequivalent parabolic fix-points (cusps). X_Γ thus extends to infinity in κ points; the latter are the Γ-orbits of the respective parabolic fix-points on $\partial\mathcal{H}$. Furthermore, X_Γ is compact iff $\kappa = 0$. In this case, the Fuchsian group Γ is called cocompact.

In the coordinates given above, the Laplace-Beltrami operator (hyperbolic Laplacian) for the Riemannian space $(\mathcal{H}, ds^2)$ reads

$$\Delta = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) . \tag{1.3}$$

On the dense domain of smooth and bounded functions in $L^2(X_\Gamma)$, the operator $-\Delta$ is essentially self-adjoint and non-negative. Henceforth, we will also denote its self-adjoint extension by $-\Delta$. Square-integrable functions on X_Γ are realised as Γ-automorphic functions $\varphi : \mathcal{H} \to \mathbb{C}$, i.e. such that $\varphi(\gamma z) = \varphi(z)$ $\forall \gamma \in \Gamma$ and

$$\|\varphi\|^2_{L^2(X_\Gamma)} = \int_{\mathcal{F}_\Gamma} |\varphi(z)|^2 \, d\mu(z) < \infty . \tag{1.4}$$

The spectral resolution of a Γ-automorphic Laplacian for a cofinite Fuchsian group $\Gamma \subset SL_2(\mathbb{R})$, strongly depends on whether Γ is cocompact or not. In the following, we want to relate the spectral resolutions of Laplacians for certain cocompact arithmetic Fuchsian groups to resolutions for non-cocompact congruence modular groups. Given an order $\mathcal{O}$ in an indefinite division quaternion algebra A over $\mathbb{Q}$, let $\mathcal{O}^1 := \{x \in \mathcal{O} : \ \mathrm{n}(x) = 1\}$ be the group of units of norm one in $\mathcal{O}$. Then $\mathcal{O}^1$ is (isomorphic to) a cocompact Fuchsian group. This gives rise to a compact orbit space $X_\mathcal{O} := \mathcal{O}^1 \backslash \mathcal{H}$. We denote the discriminant of $\mathcal{O}$ by $d = d(\mathcal{O}) \in \mathbb{N}$, and the volume of $X_\mathcal{O}$ by $A_\mathcal{O} := \mathrm{vol}(X_\mathcal{O}) < \infty$. The non-compact situation we will consider is based on the Hecke congruence groups

$$\Gamma_0(d) := \left\{ \begin{pmatrix} \alpha \, \beta \\ \gamma \, \delta \end{pmatrix} \in SL_2(\mathbb{Z}) : \ \gamma \equiv 0 \bmod d \right\} \tag{1.5}$$

of level d. These are subgroups of index

$$[\Gamma_0(1) : \Gamma_0(d)] = d \prod_{p|d} \left(1 + \frac{1}{p} \right) < \infty \tag{1.6}$$

in the full modular group $\Gamma_0(1) = SL_2(\mathbb{Z})$. Here the product extends over all prime divisors p of d. Furthermore, the groups $\Gamma_0(d)$ have

$$\kappa_d = \sum_{m|d} \phi \left((m, \frac{d}{m}) \right) \geq 1 \tag{1.7}$$

inequivalent parabolic fix-points, where ϕ denotes the Euler phi-function [16, Thm.4.2.7]. Hence, the orbit space $X_d := \Gamma_0(d) \backslash \mathcal{H}$ is non-compact, but has finite volume [16, Thms.4.1.2, 4.2.5]

$$A_d := \mathrm{vol}(X_d) = \frac{\pi}{3} d \prod_{p|d} \left(1 + \frac{1}{p} \right) . \tag{1.8}$$

The spectrum of $-\Delta$ on $L^2(X_\mathcal{O})$ is discrete, comprising of the eigenvalues

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \ldots , \qquad \lambda_n \to \infty . \tag{1.9}$$

Zero is an eigenvalue with multiplicity 1, whereas the positive eigenvalues can occur with non-trivial multiplicities. The number of eigenvalues (counted with multiplicities) $N_\mathcal{O}(\lambda) := \#\{\lambda_n \leq \lambda\}$ grows according to the Weyl asymptotics, see [6, Thm.7.1],

$$N_\mathcal{O}(\lambda) \sim \frac{A_\mathcal{O}}{4\pi} \lambda , \qquad \lambda \to \infty . \tag{1.10}$$

Let $\{\varphi_k : \ k \in \mathbb{N}_0\}$ be an orthonormal basis for $L^2(X_\mathcal{O})$ consisting of eigenfunctions of $-\Delta$, with $-\Delta \varphi_k = \lambda_k \varphi_k$, such that

$$L^2(X_\mathcal{O}) = \bigoplus_{k \in \mathbb{N}_0} [\mathbb{C} \varphi_k] . \tag{1.11}$$

Here we understand the right hand side to denote the closure of the orthogonal sum. We denote by $L^2_0(X_{\mathcal{O}})$ the closed subspace spanned by all eigenfunctions with positive eigenvalues, so that $L^2(X_{\mathcal{O}}) = [\mathbb{C}\varphi_0] \oplus L^2_0(X_{\mathcal{O}})$, where $\varphi_0(z) = (A_{\mathcal{O}})^{-\frac{1}{2}}$ is constant.

Due to the non-compactness of X_d, the spectral resolution of $-\Delta$ on $L^2(X_d)$ has a richer structure. The spectrum consists of a discrete and an absolutely continuous part; the latter is given by the interval $[\frac{1}{4}, \infty)$. The absolutely continuous subspace $\mathcal{E}_d$ of $L^2(X_d)$ is spanned by the Eisenstein series $E_s^{(i)}(z)$, $i = 1, \dots, \kappa_d$, with spectral parameter $s(1-s) \in [\frac{1}{4}, \infty)$, or $s = \frac{1}{2} + ir$, $r \in \mathbb{R}$. The constant function $g_0(z) = (A_d)^{-\frac{1}{2}}$ is an L^2-normalised eigenfunction of $-\Delta$ with eigenvalue $\mu_0 = 0$. Besides g_0, the discrete subspace of $L^2(X_d)$ is spanned by Maaß cusp forms $g : \mathcal{H} \to \mathbb{C}$, with (i) $g(\gamma z) = g(z)$ $\forall \gamma \in \Gamma_0(d)$, (ii) $-\Delta g = \mu\, g$ $(\mu > 0)$, and (iii) g vanishes at every cusp. See [6, 7, 9] for details.

For a general cofinite, non-cocompact Fuchsian group Γ, there can occur in $L^2(X_\Gamma)$ Laplace eigenfunctions which are not cusp forms. These derive from residues of Eisenstein series $E_s^{(i)}(z)$ at poles located at $s \in (\frac{1}{2}, 1]$. For the Hecke congruence groups we are considering here, the Eisenstein series have only one pole for $s \in (\frac{1}{2}, 1]$. It is located at $s = 1$, leading to the constant eigenfunction g_0, see [9, Thm.11.3].

Since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(d)$, any Laplace eigenfunction $g \in L^2(X_d)$ has to be invariant under $z \mapsto z + 1$. Hence it can be expanded in a Fourier series in $x = \operatorname{Re} z$. Cusp forms are then identified by a vanishing zero-coefficient for the Fourier expansions at every cusp. In particular, the cusp at infinity yields

$$g(z) = \sum_{n \neq 0} c(n)\, \sqrt{y}\, K_{ir}(2\pi|n|y)\, e^{2\pi i n x} \ , \tag{1.12}$$

where $\mu = r^2 + \frac{1}{4}$ is the Laplace eigenvalue corresponding to g, and $K_\nu(w)$ denotes a modified Bessel function. The cusp forms span a closed subspace $\mathcal{C}_d$ of $L^2(X_d)$, so that one has the orthogonal decomposition

$$L^2(X_d) = \mathcal{E}_d \oplus [\mathbb{C} g_0] \oplus \mathcal{C}_d \ . \tag{1.13}$$

The eigenvalues of $-\Delta$ on $[\mathbb{C} g_0] \oplus \mathcal{C}_d$,

$$0 = \mu_0 < \mu_1 \le \mu_2 \le \dots \ , \qquad \mu_n \to \infty \ , \tag{1.14}$$

follow the Weyl asymptotics [9, Ch. 11]

$$N_d(\mu) := \#\{\mu_n \le \mu\} \sim \frac{A_d}{4\pi}\, \mu \ , \qquad \mu \to \infty \ . \tag{1.15}$$

For general congruence modular groups, μ_1 is bounded from below by $\mu_1 \ge \frac{21}{100}$, see [14]. However, Selberg conjectured $\mu_1 \ge \frac{1}{4}$, see [20], so that all positive eigenvalues would be embedded in the interval $[\frac{1}{4}, \infty)$.

Given an order $\mathcal{O}$ in an indefinite quaternion division algebra over $\mathbb{Q}$, the spectral correspondence we are going to investigate is a bounded linear map from $L^2_0(X_{\mathcal{O}})$ into the space of cusp forms $\mathcal{C}_d$ for a Hecke congruence group $\Gamma_0(d)$. The level d is determined to be the discriminant $d(\mathcal{O})$ of the order $\mathcal{O}$. One thus associates to every (non-constant) eigenfunction of $-\Delta$ on $L^2(X_{\mathcal{O}})$ a Maaß cusp form in $\mathcal{C}_{d(\mathcal{O})}$, with the Laplace- and Hecke eigenvalues unchanged. This linear map is realised as an integral operator with a Siegel theta function as its integral kernel. Transformation properties of these Siegel theta functions are then the key to verify the desired properties of the spectral correspondence.

The outline of this paper is as follows: In Section 2, we recall how to calculate the quadratic Gauß sum corresponding to an arbitrary quadratic form over $\mathbb{Z}$. We then specialise in Section 3 to quadratic forms which are norm forms of quaternion algebras. If $\mathcal{O}$ is an arbitrary order in a quaternion algebra over $\mathbb{Q}$, then Theorem 3.1 gives the value of the associated quadratic Gauß sum. This is used in the proof of the transformation formula of Siegel theta functions.

The construction of the Siegel theta function entering the definition of theta-lifts is provided in Section 4. Proposition 4.1 contains the relevant transformation properties of the theta function under the unit group $\mathcal{O}^1$ as well as under the Hecke congruence group $\Gamma_0(d)$.

In Section 5, we then introduce the theta-lifts and discuss their principal properties. These lead to the conclusion in Proposition 5.4, that the Laplace eigenvalues on $L^2(X_{\mathcal{O}})$ occur (with multiplicities) among the Laplace eigenvalues on $\mathcal{C}_d$.

Specialising to Eichler orders, we study Hecke operators acting on $L^2(X_{\mathcal{O}})$ and $\mathcal{C}_d$ in Section 6. The main result in this context, stating that theta-lifts commute with the action of the Hecke operators on the two spaces, is contained in Proposition 6.1. As a consequence, Hecke eigenvalues are preserved under theta-lifts.

In Section 7, we address the problem as to whether theta lifts from $L_0^2(X_{\mathcal{O}})$ to $\mathcal{C}_d$ are newforms in $\mathcal{C}_d$. Finally, we demonstrate that whenever $\mathcal{O}$ is not a maximal order the answer is negative.

This work is a result of a cooperation that was enabled by Deutscher Akademischer Austauschdienst and Svenska Institutet. We want to thank both institutions for their support. We also would like to thank Juliusz Brzezinski for helpful discussions and valuable comments on this manuscript.

2 Quadratic Gauß sums

Let Q be an integral quadratic form on $\mathbb{Z}^n$ and A the (symmetric) matrix of the corresponding bilinear form, so that $Q(x) = \frac{1}{2}x^tAx$. Let a and c be relatively prime positive integers. In this section, we will make a general investigation of sums of the form

$$F_Q(a,c) = F_Q(\rho) := \sum_{x\in\mathbb{Z}^n/c\mathbb{Z}^n} \rho^{Q(x)}, \tag{2.1}$$

where $\rho = e^{2\pi i\frac{a}{c}}$ is a primitive c-th root of unity.

A trivial observation is that if $Q = Q_1 \perp \ldots \perp Q_m$ is an orthogonal decomposition of Q, then

$$F_Q(\rho) = \prod_{i=1}^{m} F_{Q_i}(\rho). \tag{2.2}$$

In particular, if Q is diagonal, then the calculation of $F_Q(\rho)$ simplifies to the calculation of n quadratic Gauß sums

$$G(a_i,c) := \sum_{x=1}^{c} e^{2\pi i\frac{a_i}{c}x^2}. \tag{2.3}$$

Next we will show that we may more or less reduce to the case when c is a prime power. Let p be a prime, and $c = p^sc_1$ with $(p,c_1) = 1$. If ρ is a primitive c-th root of unity, then there are primitive roots ρ_1 and ρ_p of orders c_1 and p^s respectively, such that $\rho = \rho_1\rho_p$. In this case,

$$F_Q(\rho) = \sum_{x\in\mathbb{Z}^n/c\mathbb{Z}^n} (\rho_1\rho_p)^{Q(x)} = \left(\sum_{x\in\mathbb{Z}^n/c_1\mathbb{Z}^n} \rho_1^{Q(x)}\right)\left(\sum_{x\in\mathbb{Z}^n/p^s\mathbb{Z}^n} \rho_p^{Q(x)}\right) = F_Q(\rho_1)F_Q(\rho_p), \tag{2.4}$$

where the second equality holds since $\rho_p^{Q(x)}$ only depends on x modulo $p^s\mathbb{Z}^n$. By induction, we have the following result.

Proposition 2.1 *Let $F_Q(a,c)$ be defined as in* (2.1). *If $c = \prod_p p^{s_p}$, then*

$$F_Q(a,c) = \prod_{p|c} F_Q(a_p, p^{s_p}) \tag{2.5}$$

for some integers a_p, such that $(a_p,p) = 1$.

The next step in the simplification is to prove the following.

Proposition 2.2 *Let Q and $Q^{'}$ be two integral quadratic forms, which are equivalent over the p-adic integers $\mathbb{Z}_p$. If ρ is a primitive p^s-th root of unity, then*

$$F_Q(\rho) = F_{Q'}(\rho). \tag{2.6}$$

Proof. If A and $A^{'}$ are the matrices of the corresponding bilinear forms, then Q and $Q^{'}$ being equivalent over $\mathbb{Z}_p$ means that there exists a matrix $B \in GL_n(\mathbb{Z}_p)$, such that $A^{'} = B^t AB$.

Let ρ be a primitive p^s-th root of unity. If $x \in \mathbb{Z}_p^n$, then let $\tilde{x} \in \mathbb{Z}^n$ be an arbitrary choice of a vector such that $x \equiv \tilde{x} \bmod p^s\mathbb{Z}^n$. The following reasoning will not depend on the particular choice of $\tilde{x}$. Let $y = Bx \in \mathbb{Z}^n$, with B as above. Then, when y runs through all classes modulo $p^s\mathbb{Z}^n$, so does $\tilde{x}$. Hence, we have

$$F_Q(\rho) = \sum_{y \in \mathbb{Z}^n/p^s\mathbb{Z}^n} \rho^{\frac{1}{2}y^tAy} = \sum_{Bx \in \mathbb{Z}^n/p^s\mathbb{Z}^n} \rho^{\frac{1}{2}x^t(B^tAB)x} = \sum_{\tilde{x} \in \mathbb{Z}^n/p^s\mathbb{Z}^n} \rho^{\frac{1}{2}\tilde{x}^t(B^tAB)\tilde{x}} = F_{Q'}(\rho). \tag{2.7}$$

□

Now let H and J be the quadratic forms with corresponding matrices

$$H \sim \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } J \sim \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \tag{2.8}$$

respectively. We recall the following well known result from the theory of quadratic forms. For a proof see for example [12, Ch.5].

Theorem 2.1 *Let Q be an integral quadratic form, and let p be a prime.*

1. *If p is odd, then Q is equivalent over $\mathbb{Z}_p$ to a diagonal integral quadratic form.*
2. *If $p = 2$, then Q is equivalent over $\mathbb{Z}_2$ to an orthogonal sum of a diagonal integral quadratic form and integer multiples of H and J.*

Summing up everything achieved this far, we see that calculation of $F_Q(a, c)$ has been reduced to the calculation of quadratic Gauß sums $G(a, p^s)$ for prime powers, $F_H(a, 2^s)$ and $F_J(a, 2^s)$. A proof of the following can for example be found in [13, Ch.IV.3].

Proposition 2.3 *Let p be a prime and $c = p^s$. Then*

$$G(a, c) = \begin{cases} (\frac{a}{p})^s\sqrt{c}, & \text{if } c \equiv 1 \text{ mod } 4 \\ (\frac{a}{p})i\sqrt{c}, & \text{if } c \equiv 3 \text{ mod } 4 \\ 0, & \text{if } c = 2 \\ \varphi_s(a)\sqrt{c}, & \text{if } c \equiv 0 \text{ mod } 4, \end{cases} \tag{2.9}$$

where

$$\varphi_s(a) = \begin{cases} 1+i, & \text{if } a \equiv 1 \text{ mod } 8 \\ (-1)^s(1-i), & \text{if } a \equiv 3 \text{ mod } 8 \\ (-1)^s(1+i), & \text{if } a \equiv 5 \text{ mod } 8 \\ 1-i, & \text{if } a \equiv 7 \text{ mod } 8. \end{cases} \tag{2.10}$$

We conclude this section with the calculations of $F_H(a, 2^s)$ and $F_J(a, 2^s)$.

Proposition 2.4 *Let a be an odd integer and H, J as in (2.8). If $0 \leq r < s$, then*

$$F_H(a2^r, 2^s) = 2^{r+s} \text{ and } F_J(a2^r, 2^s) = (-2)^{r+s}. \tag{2.11}$$

Proof. If $\rho_d = e^{\frac{2\pi i}{2^d}}$ and $c = 2^s$, then

$$F_H(a,c) = \sum_{x=1}^{c}\sum_{y=1}^{c} \rho_s^{axy}. \tag{2.12}$$

We now employ some well known results on sums involving roots of unity (see for example [13, Ch.IV.3]). First ay may be replaced by y, since a is odd. Furthermore, the inner sum only depends on the power of 2 in the prime factorisation of x, and is equal to 0 if x is odd. Hence, we may put $a = 1$ and $x = 2^m$. We get

$$F_H(a,c) = \sum_{m=1}^{s-1} 2^{s-m-1}\left(\sum_{y=1}^{c}\rho_s^{2^m y}\right) + \sum_{y=1}^{c}\rho_s^{2^s y} = \sum_{m=1}^{s-1} 2^{s-m-1}\left(\sum_{y=1}^{c}\rho_{s-m}^{y}\right) + c = c, \tag{2.13}$$

since the inner sum before the last equality is equal to 0. Furthermore, we have

$$F_H(a2^r, c) = \sum_{x,y=1}^{c} \rho_s^{a2^r xy} = \sum_{x,y=1}^{c} \rho_{s-r}^{axy} = 2^{2r} F_H(a, 2^{s-r}) = 2^{s+r}. \tag{2.14}$$

We now turn to F_J. It is trivial to check for $s = 1$, so assume that $s > 1$. We divide the double sum into two parts

$$F_J(a,c) = \sum_{y=1, y\, even}^{c}\left(\sum_{x=1}^{c}\rho_s^{a(x^2+xy+y^2)}\right) + \sum_{y=1, y\, odd}^{c}\rho_s^{a\frac{3}{4}y^2}\left(\sum_{x=1}^{c}\rho_s^{a(x+\frac{y}{2})^2}\right), \tag{2.15}$$

where the first part is over even y and the second one over odd y. The second part is equal to 0, since

$$\sum_{x=1}^{c}\rho_s^{a(x+\frac{1}{2})^2} = 0. \tag{2.16}$$

To prove this, we observe that $\rho_s^{a(x+\frac{1}{2})^2} = -\rho_s^{a(x+\frac{c}{2}+\frac{1}{2})^2}$. Hence, we get by Proposition 2.3 that

$$\begin{aligned} F_J(a,c) &= \frac{1}{2}\sum_{y=1, y\, even}^{2c}\left(\sum_{x=1}^{c}\rho_s^{a((x+\frac{y}{2})^2+\frac{3}{4}y^2)}\right) = \frac{1}{2}\sum_{y=1}^{c}\rho_s^{3ay^2}\left(\sum_{x=1}^{c}\rho_s^{a(x+y)^2}\right) \\ &= \frac{1}{2}G(3a,c)G(a,c) = \begin{cases} \frac{1}{2}|G(a,c)|^2 = c & \text{if } s \text{ is even,} \\ -\frac{1}{2}|G(a,c)|^2 = -c & \text{if } s \text{ is odd.} \end{cases} \end{aligned} \tag{2.17}$$

The proof for the case $r > 0$ is exactly the same as for F_H. □

3 Quaternion orders

Let A be a quaternion algebra over $\mathbb{Q}$. It is always possible to find a basis $1, j, k, jk$ of A over $\mathbb{Q}$, such that

$$j^2 = a,\ k^2 = b,\ jk = -kj,\ a, b \in \mathbb{Z} \text{ and } ab \neq 0. \tag{3.1}$$

There is a natural involution in A given by

$$x = x_0 + x_1 j + x_2 k + x_3 jk \longmapsto \overline{x} = x_0 - x_1 j - x_2 k - x_3 jk. \tag{3.2}$$

One defines the (reduced) norm, $\mathrm{n} : A \to \mathbb{Q}$, and the (reduced) trace, $\mathrm{tr} : A \to \mathbb{Q}$, by

$$\mathrm{n}(x) := x\overline{x} \text{ and } \mathrm{tr}(x) := x + \overline{x}. \tag{3.3}$$

The norm is a quadratic form on A and the corresponding bilinear form has matrix $(\mathrm{tr}(e_i\overline{e_j}))$, where $e_1,\ldots,e_4$ is a basis of A.

If $K=\mathbb{Q}_p$, the p-adic numbers, or $K=\mathbb{Q}_\infty=\mathbb{R}$, then it is well known that either $A\otimes_{\mathbb{Q}}K\cong M_2(K)$ or $A\otimes_{\mathbb{Q}}K\cong\mathbb{H}_K$, where $\mathbb{H}_K$ is a unique division algebra of dimension 4 over K. We say that A is ramified at p (at ∞), if $A\otimes_{\mathbb{Q}}\mathbb{Q}_p$ ($A\otimes_{\mathbb{Q}}\mathbb{R}$) is a division algebra. The algebra A is always ramified at an even number of places. We say that A is definite, if it is ramified at ∞; otherwise it is called indefinite. This is equivalent to the norm form being positive definite or not. The discriminant $d(A)$ of A is defined to be the product of all finite primes at which A is ramified.

An order $\mathcal{O}$ in A is a subring of A with unity which is a finitely generated $\mathbb{Z}$-module containing a basis of A. We may regard the norm on A restricted to $\mathcal{O}$ as an integral quadratic form on $\mathbb{Z}^4$, since $\mathcal{O}\cong\mathbb{Z}^4$ and $\mathrm{n}(\mathcal{O})\subseteq\mathbb{Z}$. The matrix of the bilinear form is once again given by $(\mathrm{tr}(e_i\overline{e_j}))$, where now $e_1,\ldots,e_4$ is a basis of $\mathcal{O}$ over $\mathbb{Z}$. The modulus of the determinant of this matrix is always the square of an integer, and one defines the (reduced) discriminant $d(\mathcal{O})$ of $\mathcal{O}$ to be

$$d(\mathcal{O}):=\sqrt{|\det(\mathrm{tr}(e_i\overline{e_j}))|}. \tag{3.4}$$

If $\mathcal{O}$ is a maximal order in A, then $d(A)=d(\mathcal{O})$. Moreover, if $\mathcal{O}_1\subseteq\mathcal{O}_2$, then $d(\mathcal{O}_1)=d(\mathcal{O}_2)\cdot[\mathcal{O}_2:\mathcal{O}_1]$. For further information on quaternion algebras we refer to [25].

A symmetric matrix M will be called even if $M\in M_2(\mathbb{Z})$ with even elements in the diagonal. This is obviously equivalent to the corresponding quadratic form being integral. The following lemma will be essential for the application in Section 4.

Lemma 3.1 *Let $\mathcal{O}$ be an order in a quaternion algebra over $\mathbb{Q}$, and let M be the matrix of the bilinear form corresponding to the norm form on $\mathcal{O}$. If $d=d(\mathcal{O})$, then $d\cdot M^{-1}$ is an even matrix.*

Proof. Let $e_1,\ldots,e_4$ be a basis of $\mathcal{O}$, such that $M=(\mathrm{tr}(e_i\overline{e_j}))$. The dual $\mathcal{O}^\#$ of $\mathcal{O}$ in A is $\mathcal{O}^\#:=\{x\in A:\ \mathrm{tr}(x\mathcal{O})\subseteq\mathbb{Z}\}$. The lattice $\mathcal{O}^\#$ has a basis $f_1,\ldots,f_4$, which is determined by $\mathrm{tr}(e_i\overline{f_j})=\delta_{ij}$. This basis is called the dual basis of $e_1,\ldots,e_4$. It is well known that

$$M^{-1}=(\mathrm{tr}(f_i\overline{f_j})). \tag{3.5}$$

We have, according to [3, (3.2)], that $d\cdot\mathcal{O}^\#\mathcal{O}^\#\subseteq\mathcal{O}$. Hence

$$d\cdot\mathrm{tr}(f_i\overline{f_j})=\mathrm{tr}(df_i\overline{f_j})\in\mathrm{tr}(d\cdot\mathcal{O}^\#\mathcal{O}^\#)\subseteq\mathrm{tr}(\mathcal{O})\subseteq\mathbb{Z}, \tag{3.6}$$

so that $d\cdot M^{-1}\in M_2(\mathbb{Z})$. Moreover, since $d\cdot f_i\overline{f_i}\in\mathcal{O}\cap\mathbb{Q}=\mathbb{Z}$, we get $d\cdot\mathrm{tr}(f_i\overline{f_i})\in2\mathbb{Z}$ and we are done. □

Let R be a ring. If f is a quadratic form on R^3, then $C_0(f)$ will denote the even Clifford R-algebra of f. To characterise the norm forms of quadratic orders, we will make use of a well-known one-to-one correspondence between orders in quaternion algebras and ternary quadratic forms. We will only need part of this correspondence, namely:

Lemma 3.2 *Let R be a principal ideal domain and $\mathcal{O}$ an order in a quaternion algebra over the quotient field of R. Then there is a non-degenerate R-integral quadratic form f on R^3, such that $\mathcal{O}=C_0(f)$. Furthermore, if*

$$f=\sum_{1\le i\le j\le 3}a_{ij}X_iX_j, \tag{3.7}$$

then the norm form on $C_0(f)$ is given by

$$Q=X_0^2+\sum_{(i,j,k)}\left[a_{ij}X_0X_k+a_{ii}a_{jj}X_k^2+(a_{ik}a_{jk}-a_{ij}a_{kk})X_iX_j\right], \tag{3.8}$$

where the sum is over all even permutations (i,j,k) of $(1,2,3)$.

Proof. For a proof of the first statement, see [2, (3.6)]. To prove the formula for Q, one can make use of [4, (5.2)]. We remark that in both references, the results are formulated for Gorenstein orders. However, if $\mathcal{O} = R + b(\mathcal{O})G(\mathcal{O})$, with $b(\mathcal{O})$ the Brandt invariant and $G(\mathcal{O}) \cong C_0(f)$ the Gorenstein closure (see [3] for definitions), then $\mathcal{O} \cong C_0(b(\mathcal{O}) \cdot f)$. □

Let Q be a quadratic form on L with corresponding matrix A. Then the notation

$$Q = \langle \alpha_1 \rangle \perp \ldots \perp \langle \alpha_n \rangle \tag{3.9}$$

means that there is a basis $x_1, \ldots, x_n$ of L, which is orthogonal with respect to A and satisfies $Q(x_i) = \alpha_i$.

Proposition 3.1 *Let $\mathcal{O}$ be an order in a quaternion algebra over $\mathbb{Q}$, with $d(\mathcal{O}) = \prod_p p^{d_p}$. Furthermore, let Q be the norm form on $\mathcal{O}$.*

If p is an odd prime, then there are $k, n, m, \epsilon_i \in \mathbb{N}_0$ such that

$$Q \cong \langle 1 \rangle \perp \left\langle -\epsilon_1 p^{2k} \right\rangle \perp \langle -\epsilon_2 p^n \rangle \perp \left\langle \epsilon_1\epsilon_2 p^{n+2m} \right\rangle \text{ over } \mathbb{Z}_p, \tag{3.10}$$

where $(\epsilon_i, p) = 1$. Moreover, we have $\max\{2k, n+2m\} \leq d_p = k + n + m$.

Furthermore, Q is equivalent over $\mathbb{Z}_2$ to one of the following forms:

$$\begin{aligned} Q_1 &= \langle 1 \rangle \perp \left\langle -\epsilon_1 2^{2k} \right\rangle \perp \langle -\epsilon_2 2^n \rangle \perp \left\langle \epsilon_1\epsilon_2 2^{n+2m} \right\rangle , \\ Q_2 &= H \perp 2^n H , \\ Q_3 &= \langle 1 \rangle \perp \left\langle -2^{2k-2} \right\rangle \perp 2^{k+n} H , \\ Q_4 &= J \perp 2^n J , \\ Q_5 &= \langle 1 \rangle \perp \left\langle 3 \cdot 2^{2k-2} \right\rangle \perp 2^{k+n} J , \end{aligned} \tag{3.11}$$

where H and J are defined in (2.8). In the case Q_1, ϵ_i are odd integers, $d_2 = 2 + k + n + m$ and $\max\{2k, n+2m\} \leq d_2 - 2$. Moreover, $d_2 = n$ in the cases Q_2 and Q_4, and $d_2 = 2k + n$ in the cases Q_3 and Q_5.

Proof. Let $\mathcal{O} = C_0(f)$ for a ternary integral quadratic form f.

If p is odd, then according to Theorem 2.1, we may assume that

$$f = \left\langle \delta_1 p^k \right\rangle \perp \langle \delta_2 p^n \rangle \perp \langle \delta_3 p^m \rangle , \tag{3.12}$$

where $\delta_i \in \mathbb{Z}$ and $(\delta_i, p) = 1$. From Lemma 3.2, we get that the norm form on $\mathcal{O}$ is

$$Q = \langle 1 \rangle \perp \left\langle \delta_1\delta_2 p^{k+n} \right\rangle \perp \left\langle \delta_1\delta_3 p^{k+m} \right\rangle \perp \left\langle \delta_2\delta_3 p^{n+m} \right\rangle . \tag{3.13}$$

If we rename the δ_i's and observe that at least one of the exponents has to be even, we get Q in the desired form and $\max\{k+n, k+m, n+m\} \leq k + n + m = d_p$.

If $p = 2$, then according to Theorem 2.1 we have 3 possibilities $f = f_i$ with

$$\begin{aligned} f_1 &= \left\langle \delta_1 p^k \right\rangle \perp \langle \delta_2 p^n \rangle \perp \langle \delta_3 p^m \rangle , \\ f_2 &= 2^k H \perp \langle \delta 2^n \rangle , \\ f_3 &= 2^k J \perp \langle \delta 2^n \rangle , \end{aligned} \tag{3.14}$$

where δ_i are odd. The case f_1 is completely analogous to the case with p odd. For f_2, we get from Lemma 3.2 that the norm form on $\mathcal{O}$ is equal to

$$Q = X_0^2 + 2^k X_0 X_3 - \delta 2^{k+n} X_1 X_2. \tag{3.15}$$

From this, we immediately get that $Q \cong Q_2$ if $k=0$, and $Q \cong Q_3$ if $k>0$. For f_3, we get once again from Lemma 3.2 that the norm form on $\mathcal{O}$ is equal to

$$Q = X_0^2 + 2^k X_0 X_3 + 2^{2k} X_3^2 + \delta 2^{k+n}(X_1^2 - X_1 X_2 + X_2^2). \tag{3.16}$$

This implies that $Q \cong Q_4$ if $k=0$, and $Q \cong Q_5$ if $k>0$. $\square$

Now we restrict our attention from the general situation in Section 2 to the special case, when Q is the norm form of an order in a quaternion algebra over $\mathbb{Q}$.

Theorem 3.1 *Let $\mathcal{O}$ be an order in a quaternion algebra A over $\mathbb{Q}$, and let Q be the norm form on A restricted to $\mathcal{O} \cong \mathbb{Z}^4$. Suppose that $d = d(\mathcal{O}) = \prod_p p^{d_p}$ and $c = \prod_p p^{c_p}$, with $d_p \le c_p$ for all primes p. If $(a,c)=1$, then*

$$F_Q(a,c) = \begin{cases} -c^2 d, & \text{if } A \text{ is definite,} \\ c^2 d, & \text{if } A \text{ is indefinite.} \end{cases} \tag{3.17}$$

Proof. We will prove the theorem by showing that

$$F_Q(a,p^{c_p}) = \begin{cases} -p^{2c_p+d_p}, & \text{if } A \text{ is ramified at } p, \\ p^{2c_p+d_p}, & \text{otherwise.} \end{cases} \tag{3.18}$$

The result follows from this and Proposition 2.1, since A is definite iff it is ramified at an odd number of (finite) primes.

First assume that p is odd. According to Propositions 2.2 and 3.1, we may assume that

$$Q = \langle 1 \rangle \perp \left\langle -\epsilon_1 p^{2k} \right\rangle \perp \langle -\epsilon_2 p^n \rangle \perp \left\langle \epsilon_1\epsilon_2 p^{n+2m} \right\rangle, \tag{3.19}$$

where $(\epsilon_i, p) = 1$, and $k,n,m \in \mathbb{N}_0$. We have $c_p \ge d_p = k+n+m \ge \max\{2k, n+2m\}$. If $\rho_s = e^{\frac{2\pi i}{p^s}}$ and $c = c_p$, then

$$\begin{aligned} F_Q(a,p^c) &= \left(\sum_{x_0=1}^{p^c} \rho_c^{a x_0^2}\right)\left(\sum_{x_1=1}^{p^c} \rho_c^{-a\epsilon_1 p^{2k} x_1^2}\right)\left(\sum_{x_2=1}^{p^c} \rho_c^{-a\epsilon_2 p^n x_2^2}\right)\left(\sum_{x_3=1}^{p^c} \rho_c^{a\epsilon_1\epsilon_2 p^{n+2m} x_3^2}\right) \\ &= G(a,p^c)\left(\sum_{x_1=1}^{p^c} \rho_{c-2k}^{-a\epsilon_1 x_1^2}\right)\left(\sum_{x_2=1}^{p^c} \rho_{c-n}^{-a\epsilon_2 x_2^2}\right)\left(\sum_{x_3=1}^{p^c} \rho_{c-n-2m}^{a\epsilon_1\epsilon_2 x_3^2}\right) \\ &= G(a,p^c) p^{2k} G(-a\epsilon_1, p^{c-2k}) p^n G(-a\epsilon_2, p^{c-n}) p^{n+2m} G(a\epsilon_1\epsilon_2, p^{c-n-2m}) \, . \end{aligned} \tag{3.20}$$

If we divide into different cases depending on c, n mod 2, and p mod 4, and use Proposition 2.3, we get

$$F_Q(a,p^{c_p}) = \begin{cases} p^{2c_p+d_p}, & \text{if } n \text{ is even,} \\ (\frac{\epsilon_1}{p}) p^{2c_p+d_p}, & \text{if } n \text{ is odd.} \end{cases} \tag{3.21}$$

But A is ramified at p iff n is odd and $(\frac{\epsilon_1}{p}) = -1$. To show this, one can for example calculate the Hasse invariant, $S(Q)$, of Q [12, §3.4]. Namely, we get that

$$S(Q) = (p^n, \epsilon_1)_p \, , \tag{3.22}$$

where $(a,b)_p$ is the Hilbert symbol for $\mathbb{Q}_p$. The assertion follows since A is ramified iff $S(Q) = -1$ [12, Th.3.5.1].

Now assume that $p=2$ and set $c = c_2$. This case is a little more elaborate, since we have to take some non-diagonal forms into account. By Proposition 3.1, we get 5 different possibilities $Q \cong Q_i$. Analogous to the odd case, we first get

$$\begin{aligned} F_{Q_1}(a,2^c) &= 2^{2(k+n+m)} G(a,2^c) G(-a\epsilon_1, 2^{c-2k}) G(-a\epsilon_2, 2^{c-n}) G(a\epsilon_1\epsilon_2, 2^{c-n-2m}) \\ &= 2^{2c+k+n+m} \varphi_c(a) \overline{\varphi_c(a\epsilon_1)} \varphi_{c-n}(a\epsilon_2) \varphi_{c-n}(a\epsilon_1\epsilon_2) \, , \end{aligned} \tag{3.23}$$

where φ_s is defined in Proposition 2.3. Notice that $c_2 \geq d_2 \geq \max\{2k, n+2m\}+2$ is crucial here. We have

$$\varphi_s(a) = \overline{\varphi_s(-a)} \text{ and } \varphi_s(a) = (-1)^s \overline{\varphi_s(3a)}. \tag{3.24}$$

Hence, if $a \equiv (-1)^\alpha 3^\beta \bmod 8$, then

$$F_{Q_1}(a, 2^c) = \begin{cases} F_{Q_1}(1, 2^c), & \text{if } \alpha+\beta \equiv 0 \bmod 2 \\ \overline{F_{Q_1}(1, 2^c)}, & \text{if } \alpha+\beta \equiv 1 \bmod 2, \end{cases} \tag{3.25}$$

since an even number of the integers c, $c-2k$, $c-n$ and $c-n-2m$ is odd. The calculations below will show that $F_{Q_1}(1, 2^c)$ is real, and hence $F_{Q_1}(a, 2^c) = F_{Q_1}(1, 2^c)$. To calculate $F_{Q_1}(1, 2^c)$, we once again divide into different cases depending on n mod 2. Direct calculations using Proposition 2.3 give:

1. If $n \equiv 0 \bmod 2$, then

$$F_{Q_1}(1, 2^{c_2}) = \begin{cases} 2^{2c_2+d_2}, & \text{if } \epsilon_1 \equiv 1 \vee \epsilon_2 \equiv 1 \bmod 4, \\ -2^{2c_2+d_2}, & \text{otherwise.} \end{cases} \tag{3.26}$$

2. If $n \equiv 1 \bmod 2$, then

$$F_{Q_1}(1, 2^{c_2}) = \begin{cases} 2^{2c_2+d_2}, & \text{if } (\epsilon_1 \equiv 1) \vee (\epsilon_1 \equiv 3 \wedge \epsilon_2 \equiv 3 \vee 7) \vee (\epsilon_1 \equiv 7 \wedge \epsilon_2 \equiv 1 \vee 5) \bmod 8, \\ -2^{2c_2+d_2}, & \text{otherwise.} \end{cases} \tag{3.27}$$

For the other 4 cases, we get:

$$\begin{aligned} F_{Q_2}(a, 2^c) &= F_H(a, 2^c) F_H(a2^n, 2^c) = 2^{2c+n} = 2^{2c_2+d_2} \\ F_{Q_3}(a, 2^c) &= G(a, 2^c) 2^{2k-2} \overline{G(a, 2^{c-2k+2})} F_H(a2^{k+n}, 2^c) = 2^{2c+2k+n-1} |\varphi_c(a)|^2 = 2^{2c_2+d_2} \\ F_{Q_4}(a, 2^c) &= F_J(a, 2^c) F_J(a2^n, 2^c) = (-2)^{2c+n} = (-1)^n 2^{2c_2+d_2} \\ F_{Q_5}(a, 2^c) &= G(a, 2^c) 2^{2k-2} G(3a, 2^{c-2k+2}) F_J(a2^{k+n}, 2^c) \\ &= (-1)^{n+k+c} 2^{2c+2k+n-1} \varphi_c(a) \varphi_c(3a) = (-1)^{n+k} 2^{2c_2+d_2}\ . \end{aligned} \tag{3.28}$$

To show that we have a negative sign iff $\mathcal{O}$ is in an algebra which is ramified at 2, we can for example calculate the Hasse invariant. First we observe that

$$H \cong \langle 1 \rangle \perp \langle -1 \rangle \text{ and } J \cong \langle 1 \rangle \perp \langle 3 \rangle \tag{3.29}$$

over $\mathbb{Q}_2$. If $Q = Q_1$, then the Hasse invariant $S(Q)$ is equal to

$$S(Q) = -(\epsilon_1, 2^n \epsilon_2)_2, \tag{3.30}$$

where $(a, b)_2$ is the Hilbert symbol for $\mathbb{Q}_2$. Hence, Q is non-isotropic over $\mathbb{Q}_2$ iff $(\epsilon_1, 2^n\epsilon_2)_2 = -1$ [12, Th.3.5.1]. Now it is straightforward and easy to check that the sign is the desired one in all the 5 different cases. □

4 Siegel theta functions

With the help of Theorem 3.1, we will now generalise the result in [8] from the special orders considered there to arbitrary orders and also sharpen the quantitative results. Our strategy is to define a Siegel theta function as in [8], which ensures invariance under the unit group $\mathcal{O}^1$, and then to check invariance under $\Gamma_0(d)$ along the lines of [24].

If S is a symmetric matrix in $GL_n(\mathbb{R})$, then a majorant of S is a positive definite symmetric matrix P such that $PS^{-1}P = S$. We remark that if P is a majorant of S, then B^tPB is a majorant of B^tSB for $B \in GL_n(\mathbb{R})$.

Now we fix S to be

$$S = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.1}$$

Since $S^2 = \mathrm{id}$, the identity matrix is a majorant of S. For $L_1, L_2 \in SL_2(\mathbb{R})$, we define $A(L_1, L_2) \in M_4(\mathbb{R})$ by requiring that

$$\begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \\ \delta_1 \end{pmatrix} = A(L_1, L_2) \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}, \text{ where } \begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} = L_1 \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} L_2^{-1}. \tag{4.2}$$

Since $\alpha_1\delta_1 - \beta_1\gamma_1 = \alpha\delta - \beta\gamma$, we find that $A(L_1, L_2)^t S A(L_1, L_2) = S$ and hence $A(L_1, L_2)$ is a majorant of S. For w, $z = x + iy \in \mathcal{H}$, we define

$$M_z := \begin{pmatrix} y^{\frac{1}{2}} & xy^{-\frac{1}{2}} \\ 0 & y^{-\frac{1}{2}} \end{pmatrix} \text{ and } P_{zw} := A(M_z^{-1}, M_w^{-1})^t A(M_z^{-1}, M_w^{-1}). \tag{4.3}$$

Let $\mathcal{O}$ be an order in an indefinite quaternion algebra over $\mathbb{Q}$. Then let S' be the matrix of the norm form of $\mathcal{O}$ with respect to a fixed $\mathbb{Z}$-basis. For $q \in \mathcal{O}$, let $k_q \in \mathbb{Z}^4$ be the coefficient vector of q in this basis. We fix an embedding $\sigma : \mathcal{O} \longrightarrow M_2(\mathbb{R})$. Since σ is linear, we have a unique $B \in GL_4(\mathbb{R})$ which for every q satisfies

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = Bk_q, \text{ whenever } \sigma_q := \sigma(q) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{4.4}$$

Since

$$k_q^t S' k_q = 2\mathrm{n}(q) = 2\det(\sigma_q) = 2(\alpha\delta - \beta\gamma) = (Bk_q)^t S (Bk_q), \tag{4.5}$$

we conclude that $S' = B^t S B$. For this fixed embedding of $\mathcal{O}$, we define majorants P'_{zw} of S' by $P'_{zw} := B^t P_{zw} B$.

To simplify notations, we let

$$\phi(q, z, w) := \frac{|\sigma_q \overline{w} - z|^2}{\mathrm{Im}\,(\sigma_q \overline{w})\,\mathrm{Im}\, z}, \tag{4.6}$$

where $z, w \in \mathcal{H}$ and $q \in \mathcal{O}^1$, so that $\sigma_q \in SL_2(\mathbb{R})$. We also remind of the well known identity

$$\frac{|gz - gw|^2}{\mathrm{Im}\, gz\, \mathrm{Im}\, gw} = \frac{|z - w|^2}{\mathrm{Im}\, z\, \mathrm{Im}\, w}, \tag{4.7}$$

for $z, w \in \mathbb{C} \setminus \mathbb{R}$ and $g \in SL_2(\mathbb{R})$. From this identity, we derive

$$\phi(q, \sigma_{q_1} z, w) = \phi(q_1^{-1} q, z, w), \tag{4.8}$$

for $q_1 \in \mathcal{O}^1$.

Lemma 4.1 *Let k_q, P'_{zw} and ϕ be defined as above. Then*

$$k_q^t P'_{zw} k_q = -\mathrm{n}(q)(\phi(q, z, w) + 2). \tag{4.9}$$

Proof. This is (v) of Proposition 4.1 in [8], though in the form which appears in the proof of this proposition at the end of page 138. The proof only requires elementary algebraic manipulations and (4.7). $\square$

Now fix $z_0 \in \mathcal{H}$ and let $\tau = u + iv$, $z = x + iy \in \mathcal{H}$. With $R := uS' + ivP'_{zz_0}$, we define a Siegel theta function $\theta(z;\tau)$ by

$$\theta(z;\tau) := \operatorname{Im}\tau \sum_{k\in\mathbb{Z}^4} e^{\pi i k^t R k} = \operatorname{Im}\tau \sum_{q\in\mathcal{O}} e^{\pi i k_q^t R k_q}. \tag{4.10}$$

By (4.5) and Lemma 4.1, we find that

$$\theta(z;\tau) = \operatorname{Im}\tau \sum_{q\in\mathcal{O}} e^{\pi \mathrm{n}(q)[2ui + v(\phi(q,z,z_0)+2)]}. \tag{4.11}$$

The main result of this section, which is crucial for the application in Section 5, is summarised in

Proposition 4.1 *Let $\mathcal{O}$ be an order in an indefinite quaternion algebra over $\mathbb{Q}$, with (reduced) discriminant d. Then*

1. $\theta(\sigma_q z;\tau) = \theta(z;\tau)$, $\forall q \in \mathcal{O}^1$,
2. $\theta(z;g\tau) = \theta(z;\tau)$, $\forall g \in \Gamma_0(d)$.

Proof. Take $q_1 \in \mathcal{O}$. Then by (4.8) and (4.11), we get

$$\begin{aligned}\theta(\sigma_{q_1} z;\tau) &= \operatorname{Im}\tau \sum_{q\in\mathcal{O}} e^{\pi \mathrm{n}(q)[2ui + v(\phi(q,\sigma_{q_1} z,z_0)+2)]} \\ &= \operatorname{Im}\tau \sum_{q\in\mathcal{O}} e^{\pi \mathrm{n}(q_1^{-1}q)[2ui + v(\phi(q_1^{-1}q,z,z_0)+2)]} = \theta(z;\tau),\end{aligned} \tag{4.12}$$

since q_1 is a unit in $\mathcal{O}$ with $\mathrm{n}(q_1) = 1$.

For the second part of the proof one notices that due to (4.5), S' is the matrix corresponding to the norm form n of $\mathcal{O}$. Now Lemma 3.1 states that dS'^{-1} is an even matrix, when $d = d(\mathcal{O})$ is the discriminant of $\mathcal{O}$. Therefore, according to [8, (2.7)] we obtain

$$\theta(z;g\tau) = \operatorname{Im} g\tau (\operatorname{Im}\tau)^{-1} |\gamma\tau + \delta|^2 \left[|\det S'|^{-\frac{1}{2}} \gamma^{-2} F_{\mathrm{n}}(\alpha,\gamma) \right] \theta(z;\tau) \ , \tag{4.13}$$

for any $g = \left(\begin{smallmatrix}\alpha & \beta \\ \gamma & \delta\end{smallmatrix}\right) \in \Gamma_0(d)$. A complete proof of this relation can be found in [24]. Since $d|\gamma$ and $(\alpha,\gamma) = 1$ for $g \in \Gamma_0(d)$, we can apply Theorem 3.1 in order to determine that $F_{\mathrm{n}}(\alpha,\gamma) = \gamma^2 d$. Moreover, $d = |\det S'|^{\frac{1}{2}}$ and $\operatorname{Im} g\tau = |\gamma\tau + \delta|^{-2} \operatorname{Im}\tau$, so that finally $\theta(z;g\tau) = \theta(z;\tau)$. $\square$

5 Theta-lifts

As before, let A be an indefinite quaternion division algebra over $\mathbb{Q}$ and $\mathcal{O}$ some order in A and $X_{\mathcal{O}} = \mathcal{O}^1\backslash\mathcal{H}$ the associated compact surface. Given an eigenfunction $\varphi \in L_0^2(X_{\mathcal{O}})$ with $-\Delta\varphi = \lambda\varphi$, we now consider the following linear integral transformation,

$$\Theta(\varphi)(\tau) := \int_{\mathcal{F}_{\mathcal{O}}} \theta(z;\tau)\,\varphi(z)\,d\mu(z) \ . \tag{5.1}$$

Here $z = x + iy \in \mathcal{H}$, $\tau = u + iv \in \mathcal{H}$, and $\theta(z;\tau)$ is the theta function as defined in (4.10). Consider the Hecke congruence group $\Gamma_0(d)$, with its level d given by the discriminant $d(\mathcal{O})$ of the order $\mathcal{O}$. For a cusp form $g \in \mathcal{C}_d$ with eigenvalue μ, we then introduce an associated transformation,

$$\widetilde{\Theta}(g)(z) := \int_{\mathcal{F}_d} \overline{\theta(z;\tau)}\, g(\tau)\, d\mu(\tau) \ . \tag{5.2}$$

Both $\Theta(\varphi)$ and $\widetilde{\Theta}(g)$ are known as theta-lifts of φ and g, respectively.

In order to distinguish the hyperbolic Laplacian in the z-variable from the one in the τ-variable, we temporarily employ the notations $\Delta_z = y^2(\partial_x^2 + \partial_y^2)$ and $\Delta_\tau = v^2(\partial_u^2 + \partial_v^2)$, respectively. It can be shown, either directly or by referring to [8, (12.2)], that

$$\Delta_z \theta(z;\tau) = \Delta_\tau \theta(z;\tau) \ . \tag{5.3}$$

As a consequence, we observe

Lemma 5.1 *1. If $\varphi \in L^2_0(X_{\mathcal{O}})$ is an eigenfunction of $-\Delta_z$ with eigenvalue λ, then $\Theta(\varphi)$ is an eigenfunction of $-\Delta_\tau$ with the same eigenvalue.*

2. If $g \in \mathcal{C}_d$ is an eigenfunction of $-\Delta_\tau$ with eigenvalue μ, then $\widetilde{\Theta}(g)$ is an eigenfunction of $-\Delta_z$ with the same eigenvalue.

Proof. One exploits (5.3) to obtain

$$\begin{aligned} \Delta_\tau \Theta(\varphi)(\tau) &= \int_{\mathcal{F}_{\mathcal{O}}} \Delta_\tau \theta(z;\tau)\, \varphi(z)\, d\mu(z) \\ &= \int_{\mathcal{F}_{\mathcal{O}}} \Delta_z \theta(z;\tau)\, \varphi(z)\, d\mu(z) \\ &= \int_{\mathcal{F}_{\mathcal{O}}} \left(\partial_x^2 + \partial_y^2\right) \theta(z;\tau)\, \varphi(z)\, dx\, dy \ . \end{aligned} \tag{5.4}$$

A two-fold partial integration together with the relation $-\Delta_z \varphi(z) = \lambda \varphi(z)$ then shows that

$$\begin{aligned} -\Delta_\tau \Theta(\varphi)(\tau) &= -\int_{\mathcal{F}_{\mathcal{O}}} \theta(z;\tau)\, \Delta_z \varphi(z)\, d\mu(z) \\ &= \lambda\, \Theta(\varphi)(\tau) \ . \end{aligned} \tag{5.5}$$

This completes the proof of 1. Part 2 is shown in a completely analogous manner. $\square$

According to Proposition 4.1, the theta function $\theta(z;\tau)$ is automorphic with respect to $\mathcal{O}^1$ in z and with respect to $\Gamma_0(d)$ in τ. This implies that $\Theta(\varphi)(\tau)$ is automorphic with respect to $\Gamma_0(d)$, whereas $\widetilde{\Theta}(g)(z)$ is automorphic with respect to $\mathcal{O}^1$. Therefore, $\Theta(\varphi) \in C^\infty(X_d)$ and $\widetilde{\Theta}(g) \in C^\infty(X_{\mathcal{O}})$, since both are eigenfunctions of elliptic partial differential operators. The compactness of $X_{\mathcal{O}}$ then immediately yields that $\widetilde{\Theta}(g) \in L^2_0(X_{\mathcal{O}})$. Square-integrability of $\Theta(\varphi)$ is not so easily established, but we recall from [8]:

Proposition 5.1 *Let $\varphi \in L^2_0(X_{\mathcal{O}})$ be an eigenfunction of $-\Delta_z$ with eigenvalue $\lambda = r^2 + \frac{1}{4}$. Then*

1. $\Theta(\varphi)$ has a Fourier expansion at the cusp at infinity, compare (1.12), *that reads*

$$\Theta(\varphi)(\tau) = \sum_{n \neq 0} c(n)\, \sqrt{v}\, K_{ir}(2\pi|n|v)\, e^{2\pi i n u} \ , \tag{5.6}$$

with Fourier coefficients

$$c(n) = \frac{4}{\sqrt{|n|}} \sum_{j=1}^{d(n)} \varphi\left(\gamma_j \hat{z}_0\right) \ . \tag{5.7}$$

2. $\Theta(\varphi) \in \mathcal{C}_d \subset L^2(X_d)$.

REMARKS:

1. In order to arrive at the Fourier expansion (5.6), one notices in the definition (4.10) of the theta-function that the sum over $k_q \in \mathbb{Z}^4$ may be viewed as a summation over the elements $q \in \mathcal{O}$. As in [8], this sum can be split into one over $n \in \mathbb{Z}$, and a remaining sum over the elements of the set $\mathcal{O}^n := \{q \in \mathcal{O};\ \mathrm{n}(q) = n\}$. According to the lemma on p. 118 in [5], $\mathcal{O}^n$ decomposes into a finite number of (left) cosets of the unit group $\mathcal{O}^1$,

$$\mathcal{O}^n = \bigcup_{j=1}^{d(n)} \mathcal{O}^1 \gamma_j \ . \tag{5.8}$$

This defines the quantities $d(n) \in \mathbb{N}$ and $\gamma_j \in \mathcal{O}^n$ appearing in (5.7). Moreover, $z_0 \in \mathcal{H}$ is an arbitrary reference point that enters $\Theta(\varphi)$ through the matrix P'_{zz_0} appearing in the definition (4.10) of the Siegel theta function. We have also defined $\hat{z}_0 := z_0$ for $n > 0$, and $\hat{z}_0 := \overline{z_0}$ for $n < 0$. A further discussion of the Fourier coefficients will be postponed to Section 6.

2. The Fourier expansion (5.6) also shows that the theta-lift $\Theta(\varphi)$ does not depend on the choice of the embedding $\sigma : \mathcal{O} \to M_2(\mathbb{R})$, although this appears in the Siegel theta function through the matrix B, see (4.4)–(4.10). So obviously, the theta-lift in fact only depends on the choice of the reference point z_0.

An immediate corollary that can be drawn from Proposition 5.1 is of some importance to the sequel.

Corollary 5.1 *Let φ be as in Proposition* 5.1. *Then $\Theta(\varphi) \not\equiv 0$, if the reference point $z_0 \in \mathcal{H}$ has been chosen suitably.*

Proof. Consider the term corresponding to $n = 1$ in the Fourier expansion (5.6). Since then the decomposition (5.8) is trivial, i.e. $d(1) = 1$, one can choose $\gamma_1^{(1)} = \mathrm{id}$ so that the first Fourier coefficient is given by $c(1) = 4\varphi(z_0)$. Since $\varphi \in C^\infty(X_\mathcal{O})$, $\varphi \not\equiv 0$, one can choose the reference point z_0 in such a way that $c(1) = 4\varphi(z_0) \neq 0$. This is possible for almost all $z_0 \in \mathcal{H}$ with respect to $d\mu$. With this choice, $\Theta(\varphi) \not\equiv 0$. □

By linearity, one can obviously extend the definitions of the theta-lifts Θ and $\widetilde{\Theta}$ to linear combinations of eigenfunctions. Again by linearity and due to Lemma 5.1, the resulting function is a linear combination of eigenfunctions. We now show

Proposition 5.2 *The theta-lifts can be extended to bounded linear maps*

$$\begin{aligned} \Theta &: L_0^2(X_\mathcal{O}) \to \mathcal{C}_d \ , \\ \widetilde{\Theta} &: \mathcal{C}_d \to L_0^2(X_\mathcal{O}) \ . \end{aligned} \tag{5.9}$$

Proof. We have to show the boundedness of Θ and $\widetilde{\Theta}$. To this end consider an arbitrary eigenfunction $\varphi \in L_0^2(X_\mathcal{O})$, with $-\Delta\varphi = \lambda\varphi$, $\lambda > 0$. Since φ is orthogonal in $L^2(X_\mathcal{O})$ to the constant functions, one obtains that $\int_{\mathcal{F}_\mathcal{O}} \varphi(z)\, d\mu(z) = 0$. Inserting (4.10) into (5.1) then yields

$$\Theta(\varphi)(\tau) = v \sum_{k \in \mathbb{Z}^4 \setminus \{0\}} e^{i\pi u k^t S' k} \int_{\mathcal{F}_\mathcal{O}} e^{-\pi v k^t P'_{zz_0} k}\, \varphi(z)\, d\mu(z) \ , \tag{5.10}$$

which enables the estimate

$$|\Theta(\varphi)(\tau)| \leq v \sum_{k \in \mathbb{Z}^4 \setminus \{0\}} \int_{\mathcal{F}_\mathcal{O}} e^{-\pi v k^t P'_{zz_0} k}\, |\varphi(z)|\, d\mu(z) \ . \tag{5.11}$$

Choose $w \in \mathcal{F}_{\mathcal{O}}$ such that $0 < k^t P'_{wz_0} k \leq k^t P'_{zz_0} k$ for all $k \in \mathbb{Z}^4 \backslash \{0\}$ and for all $z \in \mathcal{F}_{\mathcal{O}}$. The Hölder inequality implies

$$\|\varphi\|_{L^1(X_{\mathcal{O}})} \leq A_{\mathcal{O}} \; \|\varphi\|_{L^2(X_{\mathcal{O}})} \tag{5.12}$$

so that finally, upon squaring (5.11) and integrating the result over $\mathcal{F}_d$,

$$\begin{aligned} \|\Theta(\varphi)\|^2_{L^2(X_d)} &\leq A^2_{\mathcal{O}} \sum_{k,l \in \mathbb{Z}^4 \backslash \{0\}} \int_{\mathcal{F}_d} e^{-\pi v (k^t P'_{wz_0} k + l^t P'_{wz_0} l)} \, du \, dv \;\; \|\varphi\|^2_{L^2(X_{\mathcal{O}})} \\ &= C^2 \;\; \|\varphi\|^2_{L^2(X_{\mathcal{O}})} \;, \end{aligned} \tag{5.13}$$

with some constant $C > 0$.

Now let $\varphi \in L^2_0(X_{\mathcal{O}})$ be arbitrary. Denote the distinct eigenvalues of the Laplacian on $L^2_0(X_{\mathcal{O}})$ by $0 < \tilde{\lambda}_1 < \tilde{\lambda}_2 < \ldots$, and the respective eigenspaces by $\mathrm{Eig}(\tilde{\lambda}_k) \subset L^2_0(X_{\mathcal{O}})$; the latter subspaces are clearly mutually orthogonal. Let $\tilde{\varphi}_k$ be the projection of φ onto $\mathrm{Eig}(\tilde{\lambda}_k)$ so that $\varphi = \sum_k \tilde{\varphi}_k$. By linearity of Θ, then

$$\Theta(\varphi) = \sum_k \Theta(\tilde{\varphi}_k) \; . \tag{5.14}$$

Since $\Theta(\tilde{\varphi}_k)$ and $\Theta(\tilde{\varphi}_l)$ are eigenfunctions of the Laplacian with eigenvalues $\tilde{\lambda}_k \neq \tilde{\lambda}_l$ for $k \neq l$, the summands appearing on the r.h.s. of (5.14) are mutually orthogonal. Hence

$$\|\Theta(\varphi)\|^2_{L^2(X_d)} = \sum_k \|\Theta(\tilde{\varphi}_k)\|^2_{L^2(X_d)} \leq C^2 \sum_k \|\tilde{\varphi}_k\|^2_{L^2(X_{\mathcal{O}})} = C^2 \;\; \|\varphi\|^2_{L^2(X_{\mathcal{O}})} \;, \tag{5.15}$$

and thus Θ is bounded. The proof of the boundedness of $\widetilde{\Theta}$ is completely analogous. □

Proposition 5.1 in conjunction with Corollary 5.1 shows that an eigenvalue λ of $-\Delta$ on $L^2(X_{\mathcal{O}})$ necessarily also occurs in the discrete spectrum of $-\Delta$ on $L^2(X_d)$, since the eigenfunction $\varphi \in L^2(X_{\mathcal{O}})$, $-\Delta\varphi = \lambda\varphi$, is mapped to a not identically vanishing eigenfunction $\Theta(\varphi) \in L^2(X_d)$ with eigenvalue λ. However, if Θ is not injective, then multiplicities will be reduced. We thus have to have a closer look at the properties of the linear maps Θ and $\widetilde{\Theta}$, and describe their ranges and kernels as far as possible. To this end, we introduce the closed linear subspaces $V \subset \mathcal{C}_d$ and $W \subset L^2_0(X_{\mathcal{O}})$ as the orthogonal complements of $\mathrm{ran}\,\Theta$ and $\mathrm{ran}\,\widetilde{\Theta}$, respectively. That is,

$$\begin{aligned} \mathcal{C}_d &= \mathrm{ran}\,\Theta \oplus V \;, \\ L^2_0(X_{\mathcal{O}}) &= \mathrm{ran}\,\widetilde{\Theta} \oplus W \;. \end{aligned} \tag{5.16}$$

The r.h.s.'s of (5.16) are to be understood as closures of direct sums, which are orthogonal with respect to the scalar products $\langle \cdot, \cdot \rangle_{X_{\mathcal{O}}}$ and $\langle \cdot, \cdot \rangle_{X_d}$ in $L^2(X_{\mathcal{O}})$ and $L^2(X_d)$, respectively. In order to characterise V and W further, the following observation will prove useful.

Lemma 5.2 *The maps Θ and $\widetilde{\Theta}$ are mutually adjoint, that is if $\varphi \in L^2_0(X_{\mathcal{O}})$ and $g \in \mathcal{C}_d$, then*

$$\langle \Theta(\varphi), g \rangle_{X_d} = \left\langle \varphi, \widetilde{\Theta}(g) \right\rangle_{X_{\mathcal{O}}} \; . \tag{5.17}$$

Proof. One simply inserts the definition (5.1) into the l.h.s. of (5.17) and observes

$$\begin{aligned} \langle \Theta(\varphi), g \rangle_{X_d} &= \int_{\mathcal{F}_d} \overline{\int_{\mathcal{F}_{\mathcal{O}}} \theta(z;\tau)\, \varphi(z)\, d\mu(z)} \; g(\tau)\, d\mu(\tau) \\ &= \int_{\mathcal{F}_{\mathcal{O}}} \overline{\varphi(z)} \int_{\mathcal{F}_d} \overline{\theta(z;\tau)}\, g(\tau)\, d\mu(\tau)\, d\mu(z) \\ &= \left\langle \varphi, \widetilde{\Theta}(g) \right\rangle_{X_{\mathcal{O}}} \; . \end{aligned} \tag{5.18}$$

□

This fact results in a further important observation.

Proposition 5.3 *1.* $\mathcal{C}_d = ran\,\Theta \oplus ker\,\widetilde{\Theta}$,

2. $L_0^2(X_{\mathcal{O}}) = ran\,\widetilde{\Theta}$,

3. $\Theta : L_0^2(X_{\mathcal{O}}) \to \mathcal{C}_d$ *is injective.*

Proof. In order to prove 1. and 2., we first observe that the decompositions $\mathcal{C}_d = \operatorname{ran}\Theta \oplus \ker\widetilde{\Theta}$ and $L_0^2(X_{\mathcal{O}}) = \operatorname{ran}\widetilde{\Theta} \oplus \ker\Theta$ follow from Θ and $\widetilde{\Theta}$ being mutually adjoint: take any $g \in V = (\operatorname{ran}\Theta)^{\perp}$ and an arbitrary $\varphi \in L_0^2(X_{\mathcal{O}})$. The equality

$$0 = \langle \Theta(\varphi), g \rangle_{X_d} = \left\langle \varphi, \widetilde{\Theta}(g) \right\rangle_{X_{\mathcal{O}}} \tag{5.19}$$

then implies that $V \subseteq \ker\widetilde{\Theta}$. Conversely, let $g \in \ker\widetilde{\Theta}$ and $\varphi \in L_0^2(X_{\mathcal{O}})$. Then

$$0 = \left\langle \varphi, \widetilde{\Theta}(g) \right\rangle_{X_{\mathcal{O}}} = \langle \Theta(\varphi), g \rangle_{X_d} \tag{5.20}$$

shows that $\ker\widetilde{\Theta} \subseteq (\operatorname{ran}\Theta)^{\perp} = V$, and hence $V = \ker\widetilde{\Theta}$. Interchanging Θ and $\widetilde{\Theta}$ yields $W = \ker\Theta$.

To complete the proof of 2., we have to show that $W = \ker\Theta = \{0\}$. To this end take any $\varphi \in \ker\Theta$, so that $\Theta(\varphi) = 0$ a.e. Decompose φ into its projections $\tilde{\varphi}_k$ onto $\operatorname{Eig}(\tilde{\lambda}_k)$ as in (5.14). Then

$$0 = \Theta(\varphi) = \sum_k \Theta(\tilde{\varphi}_k) \quad \text{a.e.} \tag{5.21}$$

Due to the mutual orthogonality of the summands appearing on the r.h.s. of (5.21) these are linearly independent. However, this contradicts (5.21) unless $\Theta(\tilde{\varphi}_k) = 0$ a.e. for all k. Corollary 5.1 then implies that $\tilde{\varphi}_k = 0$ a.e. for all k, so that finally $\varphi = 0$ a.e. Since $\ker\Theta = \{0\}$, the last assertion is obvious. □

We now restrict $\widetilde{\Theta} : \mathcal{C}_d = \operatorname{ran}\Theta \oplus \ker\widetilde{\Theta} \to L_0^2(X_{\mathcal{O}})$ to the complement of its kernel, so that we have the two vector space isomorphisms Θ and $\widetilde{\Theta}|_{\operatorname{ran}\Theta}$ between $L_0^2(X_{\mathcal{O}})$ and $\operatorname{ran}\Theta \subset \mathcal{C}_d$. Moreover, both isomorphisms respect the spectral decompositions of $L_0^2(X_{\mathcal{O}})$ and $\mathcal{C}_d$, respectively. This observation immediately yields

Proposition 5.4 *The (discrete) spectrum of $-\Delta$ on $L^2(X_{\mathcal{O}})$, including multiplicities, is completely contained in the discrete spectrum of $-\Delta$ on $L^2(X_d)$.*

As a consequence one for example observes that the spectral counting functions $N_{\mathcal{O}}(\lambda)$, see (1.10), and $N_d(\mu)$, see (1.15), obey

$$N_{\mathcal{O}}(\lambda) \le N_d(\lambda) \ . \tag{5.22}$$

The Weyl asymptotics (1.10) and (1.15) therefore allow to obtain the geometric information $A_{\mathcal{O}} \le A_d$ from the spectral information contained in Proposition 5.4 through (5.22).

However, one can learn even more about the geometry of compact surfaces of the type $X_{\mathcal{O}}$ from the above spectral correspondence. To this end one employs Proposition 5.4 to obtain the lower bound $\lambda_1 \ge \frac{21}{100}$ for the smallest positive eigenvalue of $-\Delta$ on $L^2(X_{\mathcal{O}})$ from the respective bound $\mu_1 \ge \frac{21}{100}$ for $\Gamma_0(d)$ [14]. We then exploit a relation between short closed geodesics on closed compact surfaces $X_\Gamma = \Gamma\backslash\mathcal{H}$ (of genus $g \ge 2$) and small eigenvalues of the Laplacian on $L^2(X_\Gamma)$. Denote by L_{X_Γ} the length of the shortest dissecting closed geodesic on X_Γ, i.e. the shortest geodesic in the zero class in $H_1(X_\Gamma, \mathbb{Z})$. In [19] one then finds

Proposition 5.5 *There exist constants $c_1(g) > 0$ and $c_2(g) > 0$ depending only on the genus, such that for any closed compact surface X_Γ of genus g,*

$$c_1(g)\, L_{X_\Gamma} \leq \lambda_1(X_\Gamma) \leq c_2(g)\, L_{X_\Gamma} \ , \tag{5.23}$$

where $\lambda_1(X_\Gamma)$ is the smallest positive eigenvalue of $-\Delta$ on $L^2(X_\Gamma)$.

Any order $\mathcal{O}$ in an indefinite division quaternion algebra over $\mathbb{Q}$ with a unit group $\mathcal{O}^1$ void of elliptic elements yields a closed and compact surface $X_\mathcal{O}$ of some genus $g \geq 2$. The bound $\lambda_1(X_\mathcal{O}) \geq \frac{21}{100}$ then implies that $L_{X_\mathcal{O}} \geq \frac{21}{100}\frac{1}{c_2(g)}$. Since $c_2(g)$ only depends on the genus, one has thus obtained a lower bound for the shortest dissecting closed geodesic on $X_\mathcal{O}$. A surface of the type $X_\mathcal{O}$ therefore cannot develop a too thin neck.

6 Hecke operators

The arithmetic nature of the surfaces $X_\mathcal{O}$ and X_d allows one to introduce non-trivial correspondences on them. These then give rise to Hecke operators acting on $L^2(X_\mathcal{O})$ and $L^2(X_d)$, respectively.

In the cocompact case we restrict our attention to Eichler orders $\mathcal{O}$ of level N in an indefinite division quaternion algebra A over $\mathbb{Q}$, see [16, §5.3] for explanations. Here we only remark that N has to be coprime to the discriminant $d(A)$ of A, i.e. N is not divisible by any ramified prime of A. Furthermore, the discriminant $d = d(\mathcal{O})$ of the order is given by the product $d(\mathcal{O}) = Nd(A)$. Hence maximal orders are characterised as Eichler orders of level one. The reason for this restriction is that the Hecke operators are not known explicitly in general.

In order to define Hecke operators, we follow the general prescription, as for example outlined in [16, §2.7]. One starts with an arbitrary element $u \neq 0$ of the commensurator of $\mathcal{O}^1$, which in our situation is given by A itself [25, Ch.IV,Prop.1.4]. Then $\mathcal{O}(u) := \mathcal{O}^1 \cap u^{-1}\mathcal{O}^1 u$ is of finite index $d(u)$ in $\mathcal{O}^1$, so that

$$\mathcal{O}^1 = \bigcup_{j=1}^{d(u)} \mathcal{O}(u)\,\varepsilon_j \ , \tag{6.1}$$

for some representatives $\varepsilon_1, \dots, \varepsilon_{d(u)} \in \mathcal{O}^1$. Since the norm $\mathrm{n}(u)$ is rational, one can choose $q \in \mathbb{Z}$ such that $\mathrm{n}(qu) = q^2\mathrm{n}(u) \in \mathbb{Z}$. However, $\mathcal{O}(qu) = \mathcal{O}(u)$ so that we can restrict to $u \in A$ with integral norm. Moreover, an Eichler order always contains an element ε with norm $\mathrm{n}(\varepsilon) = -1$, see [11, Prop.6.2,Cor.2]. This allows to restrict ones attention to $\mathrm{n}(u) \geq 1$, since if $\mathrm{n}(u) < 0$ one can change to εu with norm $\mathrm{n}(\varepsilon u) = -\mathrm{n}(u)$ and $\mathcal{O}(\varepsilon u) = \mathcal{O}(u)$.

The decomposition (6.1) of the proper unit group now yields a corresponding decomposition

$$\mathcal{O}^1 u \mathcal{O}^1 = \bigcup_{j=1}^{d(u)} \mathcal{O}^1 u \varepsilon_j \tag{6.2}$$

of the double coset $\mathcal{O}^1 u \mathcal{O}^1$. Following the general scheme, one can give the set $\mathcal{R}(\mathcal{O}) := \{\mathcal{O}^1 u \mathcal{O}^1;\ u \in \mathcal{O},\ \mathrm{n}(u) \geq 1\}$ a ring structure, see [16, §2.7]. $\mathcal{R}(\mathcal{O})$ is called the Hecke ring of $\mathcal{O}$ and in case of an Eichler order is known to be commutative, see [16, Cor.5.3.7].

The Hecke ring $\mathcal{R}(\mathcal{O})$ can be represented on $L^2(X_\mathcal{O})$ by introducing operators $\hat{T}_u$ through

$$\left(\hat{T}_u \varphi\right)(z) := \frac{1}{\sqrt{|\mathrm{n}(u)|}} \sum_{j=1}^{d(u)} \varphi(u\varepsilon_j z) \ . \tag{6.3}$$

One then defines for $n \geq 1$ the Hecke operators

$$\widetilde{T}_n := \sum_{u \in \mathcal{O}^n} \hat{T}_u \ , \tag{6.4}$$

where the sum is over representatives $u \in \mathcal{O}^n$ that yield distinct double cosets $\mathcal{O}^1 u \mathcal{O}^1$. Returning to the decomposition (5.8) of $\mathcal{O}^n$ and noticing that

$$\mathcal{O}^n = \bigcup_{u \in \mathcal{O}^n} \mathcal{O}^1 u \mathcal{O}^1 \ , \tag{6.5}$$

we obtain that the Hecke operator (6.4) applied to $\varphi \in L^2(X_{\mathcal{O}})$ reads

$$\left(\widetilde{T}_n \varphi\right)(z) = \frac{1}{\sqrt{n}} \sum_{j=1}^{d(n)} \varphi(\gamma_j z) \ . \tag{6.6}$$

If in (5.8) $n = \mathrm{n}(u)$ is negative, we choose $\varepsilon \in \mathcal{O}$ with norm $\mathrm{n}(\varepsilon) = -1$ and consider εu with norm $\mathrm{n}(\varepsilon u) = -n \geq 1$. The corresponding Hecke operator is then

$$\left(\widetilde{T}_n \varphi\right)(z) = \frac{1}{\sqrt{|n|}} \sum_{j=1}^{d(n)} \varphi(\gamma_j z) = \left(\widetilde{T}_{-n} \varphi\right)(z) \ . \tag{6.7}$$

It is well known, and can be readily deduced from (6.6), that the $\widetilde{T}_n$ are bounded linear operators on $L^2(X_{\mathcal{O}})$ which commute with the hyperbolic Laplacian. They are moreover self-adjoint, a fact that can be drawn from the observation that if $u \in \mathcal{O}^n$ defines the operator $\widetilde{T}_n$, then the adjoint operator $\widetilde{T}_n^*$ is generated by the conjugate $\overline{u}$ of u,

$$\begin{aligned}
\left\langle \widetilde{T}_n \varphi, \psi \right\rangle_{X_{\mathcal{O}}} &= \frac{1}{\sqrt{|n|}} \sum_{j=1}^{d(n)} \int_{\mathcal{F}_{\mathcal{O}}} \overline{\varphi(u\varepsilon_j z)}\, \psi(z)\, d\mu(z) \\
&= \frac{1}{\sqrt{|n|}} \sum_{j=1}^{d(n)} \int_{\mathcal{F}'_{\mathcal{O}}} \overline{\varphi(w)}\, \psi(u^{-1}\varepsilon_j w)\, d\mu(w) \\
&= \left\langle \varphi, \widetilde{T}_n^* \psi \right\rangle_{X_{\mathcal{O}}} \ .
\end{aligned} \tag{6.8}$$

Since $u^{-1} = \frac{1}{n}\overline{u}$, so that $u^{-1}z = \overline{u}z$, the assertion follows from the boundedness of $\widetilde{T}_n$. However, $\mathrm{n}(\overline{u}) = \mathrm{n}(u) = n$ implies that $\widetilde{T}_n^* = \widetilde{T}_n$ since the Hecke operators obviously only depend on n.

The Hecke operators $\widetilde{T}_n$ for $n \in \mathbb{N}$ commute, since the Hecke ring $\mathcal{R}(\mathcal{O})$ is commutative. Hence, one can choose an orthonormal basis $\{\varphi_k;\ k \in \mathbb{N}_0\}$ of $L^2(X_{\mathcal{O}})$ consisting of simultaneous eigenfunctions of the Laplacian and the Hecke operators,

$$\begin{aligned}
-\Delta \varphi_k &= \lambda_k \varphi_k \ , \\
\widetilde{T}_n \varphi_k &= \tilde{t}_k(n) \varphi_k \ .
\end{aligned} \tag{6.9}$$

Such an orthonormal system is called a Hecke basis of $L^2(X_{\mathcal{O}})$.

For the sequel, we will need the multiplicative properties of the Hecke operators. These can be derived immediately from the multiplication rules in the Hecke ring $\mathcal{R}(\mathcal{O})$ as given in [16, Cor.5.3.7]. First notice that the elements $T(p,p) \in \mathcal{R}(\mathcal{O})$, $(p, d(\mathcal{O})) = 1$, as they appear in [16, Cor.5.3.7], are represented on $L^2(X_{\mathcal{O}})$ by p^{-1} times the identity. This leads to the observations that

$$\widetilde{T}_n \widetilde{T}_m = \widetilde{T}_{nm} \ , \quad \text{if } (n,m) = 1 \ , \tag{6.10}$$

$$\widetilde{T}_p \widetilde{T}_{p^e} = \begin{cases} \widetilde{T}_{p^{e+1}} + \widetilde{T}_{p^{e-1}}, & \text{if } (p, d(\mathcal{O})) = 1, \\ \widetilde{T}_{p^{e+1}}, & \text{if } (p, d(\mathcal{O})) > 1, \end{cases} \tag{6.11}$$

for all $m, n \in \mathbb{N}$ and primes p. This enables us to obtain

Lemma 6.1 *The Hecke operators $\widetilde{T}_n$, $n \in \mathbb{N}$, on $L^2(X_{\mathcal{O}})$ obey*

$$\widetilde{T}_n \widetilde{T}_m = \sum_{\substack{l|(n,m) \\ (l,d(\mathcal{O}))=1}} \widetilde{T}_{nm/l^2} \ . \tag{6.12}$$

Proof. By induction, one can derive from (6.11) that

$$\widetilde{T}_{p^f} \widetilde{T}_{p^e} = \begin{cases} \sum_{i=0}^{\min\{e,f\}} \widetilde{T}_{p^{e+f-2i}}, & \text{if } (p, d(\mathcal{O})) = 1, \\ \widetilde{T}_{p^{e+f}}, & \text{if } (p, d(\mathcal{O})) > 1. \end{cases} \tag{6.13}$$

Together with (6.10) one immediately gets (6.12). □

If one now considers a Hecke basis $\{\varphi_k;\ k \in \mathbb{N}_0\}$, Lemma 6.1 readily yields the multiplicative properties of the Hecke eigenvalues $\tilde{t}_k(n)$ as

$$\tilde{t}_k(n)\, \tilde{t}_k(m) = \sum_{\substack{l|(n,m) \\ (l,d(\mathcal{O}))=1}} \tilde{t}_k \left(\frac{nm}{l^2} \right) \tag{6.14}$$

for all $n, m \in \mathbb{N}$, $k \in \mathbb{N}_0$.

In the case of the congruence modular group $\Gamma_0(d)$, the construction of its Hecke ring $\mathcal{R}(\Gamma_0(d))$ is based on the same principles as for the unit group $\mathcal{O}^1$ discussed above. The result is well known, and the procedure is for example explained in [16, §4.5]. If $n \in \mathbb{N}$ with $(n, d) = 1$, the action of the Hecke operator T_n on $g \in \mathcal{C}_d$ is the same as in the case of the full modular group,

$$(T_n g)\,(\tau) = \frac{1}{\sqrt{n}} \sum_{\substack{\alpha\delta=n \\ 0\leq\beta<\delta}} g \left(\frac{\alpha\tau + \beta}{\delta} \right) \ . \tag{6.15}$$

Thus the Hecke operators T_n, $(n, d) = 1$, together with the hyperbolic Laplacian form a commutative ring of self-adjoint operators. One can therefore introduce a Hecke basis $\{g_k;\ k \in \mathbb{N}\}$ for $\mathcal{C}_d$,

$$\begin{aligned} -\Delta g_k &= \mu_k g_k \ , \\ T_n g_k &= t_k(n) g_k \ , \end{aligned} \tag{6.16}$$

where $(n, d) = 1$. As in (1.12), every eigenform g_k admits a Fourier expansion

$$g_k(\tau) = \sum_{n\neq 0} c_k(n) \sqrt{v}\, K_{ir}(2\pi|n|v)\, e^{2\pi i n u} \ . \tag{6.17}$$

In complete analogy to [16, Lem.4.5.15] one obtains for the Fourier coefficients of g_k

Lemma 6.2 *The Fourier coefficients $c_k(n)$ with $(n, d) = 1$ obey $c_k(n) = c_k(1) t_k(n)$.*

Hence, all Fourier coefficients $c_k(n)$ with $(n, d) = 1$ vanish whenever $c_k(1) = 0$. On the other hand, if $c_k(1) \neq 0$, one can express the Hecke eigenvalues in terms of the Fourier coefficients as $t_k(n) = \frac{c_k(n)}{c_k(1)}$, $(n, d) = 1$.

To complete the discussion of Hecke operators on $\mathcal{C}_d$, one still has to introduce Hecke operators T_p for primes p dividing the level d. In explicit terms these read

$$(T_p g)\,(\tau) = \frac{1}{\sqrt{p}} \sum_{\beta=0}^{p-1} g \left(\frac{\tau + \beta}{p} \right) \ , \tag{6.18}$$

compare [16, Lem.4.5.6]. The multiplication rules of the Hecke operators on $\mathcal{C}_d$ can now be summarised in

$$\begin{aligned} T_n\,T_m &= T_{nm}, \quad \text{if } (n,m)=1,\\ T_{p^e} &= (T_p)^e\,, \quad \text{if } p \text{ prime and } p|d, \end{aligned}$$

compare [16, Thm4.5.13] and [9, (8.39)].

Our primary goal in this section is to study the composition of Hecke operators and theta-lifts. This discussion is based on Proposition 5.1, which yields the Fourier coefficients of a theta-lift of a Laplace eigenfunction $\varphi \in L_0^2(X_{\mathcal{O}})$. We stress that in this context, it is essential to choose the level d of the congruence modular group $\Gamma_0(d)$ to be the discriminant $d(\mathcal{O})$ of the quaternion order $\mathcal{O}$.

Now let $\varphi \in L_0^2(X_{\mathcal{O}})$ be an eigenfunction of the Laplacian, $-\Delta\varphi = \lambda\varphi$. According to Proposition 5.1, (6.6) and (6.7), the Fourier expansion of the theta-lift reads

$$\Theta(\varphi)(\tau) = \sum_{n\neq 0} 4\left(\widetilde{T}_n\varphi\right)(\hat{z}_0)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\,. \tag{6.19}$$

We now consider a Hecke basis $\{\varphi_k;\ k\in\mathbb{N}\}$ of $L_0^2(X_{\mathcal{O}})$, so that

$$\begin{aligned} \Theta(\varphi_k)(\tau) &= 4\varphi_k(z_0)\sum_{n>0}\tilde{t}_k(n)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\\ &\quad +4\varphi_k(\overline{z_0})\sum_{n<0}\tilde{t}_k(n)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\\ &= 4\sum_{n=1}^{\infty}\tilde{t}_k(n)\,\sqrt{v}\,K_{ir}(2\pi n v)\,\left[\varphi_k(z_0)\,e^{2\pi i n u}+\varphi_k(\overline{z_0})\,e^{-2\pi i n u}\right]\,. \end{aligned} \tag{6.20}$$

Hence, in case $\varphi_k(z_0) = \pm\varphi_k(\overline{z_0})$ the theta-lift can be expressed in terms of a cosine- or sine-Fourier series, respectively. In Corollary 5.1, we saw that generically the first Fourier coefficient $c(1)$ of a theta-lift does not vanish. In (6.20) we now explicitly see what happens if the reference point z_0 is chosen such that $\varphi_k(z_0) = 0 = \varphi_k(\overline{z_0})$: in this case $\Theta(\varphi_k)\equiv 0$.

We are now able to prove the main result of this section.

Proposition 6.1 *Let $n\in\mathbb{N}$ be arbitrary. Then*

1. *For $\varphi \in L_0^2(X_{\mathcal{O}})$ one obtains*
$$\Theta\left(\widetilde{T}_n\varphi\right) = T_n\left(\Theta\left(\varphi\right)\right)\,. \tag{6.21}$$

2. *For $g \in ran\,\Theta \subset \mathcal{C}_d$ one obtains*
$$\widetilde{\Theta}\left(T_n g\right) = \widetilde{T}_n\left(\widetilde{\Theta}\left(g\right)\right)\,. \tag{6.22}$$

Proof. First let $\varphi \in L_0^2(X_{\mathcal{O}})$ be a Laplace eigenfunction, and $(m, d(\mathcal{O})) = 1$. Then in view of (6.19) and Lemma 6.1, one finds

$$\begin{aligned} \Theta(\widetilde{T}_m\varphi)(\tau) &= \sum_{n\neq 0} 4\left(\widetilde{T}_n\widetilde{T}_m\varphi\right)(\hat{z}_0)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\\ &= \sum_{n\neq 0} 4 \sum_{\substack{l|(n,m)\\ (l,d(\mathcal{O}))=1}}\left(\widetilde{T}_{nm/l^2}\varphi\right)(\hat{z}_0)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\,. \end{aligned} \tag{6.23}$$

On the other hand, according to the well known action of T_m, $(m,d)=1$, on $\mathcal{C}_d$,

$$T_m\left(\Theta\left(\varphi\right)\right)(\tau) = \sum_{n\neq 0} 4 \sum_{l|(n,m)}\left(\widetilde{T}_{nm/l^2}\varphi\right)(\hat{z}_0)\,\sqrt{v}\,K_{ir}(2\pi|n|v)\,e^{2\pi i n u}\,. \tag{6.24}$$

Now the r.h.s.'s of (6.23) and (6.24) coincide, since $(m, d(\mathcal{O})) = 1$ and $l|(n,m)$ automatically implies that $(l, d(\mathcal{O})) = 1$.

For the case of $m = p$ a prime dividing d, we first recall that

$$\frac{1}{p}\sum_{\beta=0}^{p-1} e^{2\pi i \frac{n}{p}\beta} = \begin{cases} 1, & \text{if } p|n, \\ 0, & \text{otherwise.} \end{cases} \tag{6.25}$$

Employing the definition (6.18) of T_p for $p|d$ and Lemma 6.1, we find

$$\begin{aligned} T_p\left(\Theta(\varphi)\right)(\tau) &= \frac{1}{\sqrt{p}}\sum_{\beta=0}^{p-1}\Theta(\varphi)\left(\frac{\tau+\beta}{p}\right) \\ &= \sum_{n\neq 0} 4\left(\widetilde{T}_n\varphi\right)(\hat{z}_0)\sqrt{v}\;K_{ir}\left(2\pi|n|\frac{v}{p}\right)\,e^{2\pi i\frac{n}{p}u}\,\frac{1}{p}\sum_{\beta=0}^{p-1}e^{2\pi i\frac{n}{p}\beta} \\ &= \sum_{m\neq 0} 4\left(\widetilde{T}_{mp}\varphi\right)(\hat{z}_0)\sqrt{v}\,K_{ir}(2\pi|m|v)\,e^{2\pi imu} \\ &= \sum_{m\neq 0} 4\left(\widetilde{T}_m\widetilde{T}_p\varphi\right)(\hat{z}_0)\sqrt{v}\,K_{ir}(2\pi|m|v)\,e^{2\pi imu} \\ &= \Theta\left(\widetilde{T}_p\varphi\right)(\tau)\;. \end{aligned} \tag{6.26}$$

We have thus proven (6.21) for n being an arbitrary prime power, $n = p^e$. In fact, the powers p^e of primes $(p, d(\mathcal{O})) = 1$ are covered by (6.23)–(6.24). The case of prime powers of p dividing $d(\mathcal{O})$ then follows from (6.26), since $\widetilde{T}_{p^e} = \left(\widetilde{T}_p\right)^e$ and $T_{p^e} = (T_p)^e$. Moreover, for coprime $n, m \in \mathbb{N}$ the Hecke operators are multiplicative, as seen in (6.10) and (6). If then $n = \prod p^{e_p}$ is the prime factorisation of an arbitrary positive integer n, one finds

$$\widetilde{T}_n = \prod_p \widetilde{T}_{p^{e_p}} \quad \text{and} \quad T_n = \prod_p T_{p^{e_p}}\;. \tag{6.27}$$

Exploiting this, one applies the relation (6.21) for the different prime powers successively.

By choosing in $L_0^2(X_\mathcal{O})$ a basis of Laplace eigenfunctions, and by the linearity of the theta-lifts and the Hecke operators, respectively, one extends the validity of (6.21) to arbitrary $\varphi \in L_0^2(X_\mathcal{O})$.

The relation (6.22) follows from (6.21) by an application of Lemma 5.2. Since if $g \in \mathcal{C}_d$ and $\varphi \in L^2(X_\mathcal{O})$, then

$$\begin{aligned} \left\langle\widetilde{\Theta}\left(T_n g\right), \varphi\right\rangle_{X_\mathcal{O}} &= \left\langle T_n g, \Theta\left(\varphi\right)\right\rangle_{X_d} = \left\langle g, T_n\left(\Theta(\varphi)\right)\right\rangle_{X_d} \\ &= \left\langle g, \Theta\left(\widetilde{T}_n\varphi\right)\right\rangle_{X_d} = \left\langle\widetilde{\Theta}\left(g\right), \widetilde{T}_n\varphi\right\rangle_{X_\mathcal{O}} \\ &= \left\langle\widetilde{T}_n\left(\widetilde{\Theta}(g)\right), \varphi\right\rangle_{X_\mathcal{O}}\;. \end{aligned} \tag{6.28}$$

which finishes the proof, since φ can be chosen arbitrarily. □

If one now considers a Hecke basis $\{\varphi_k; k \in \mathbb{N}\}$ of $L_0^2(X_\mathcal{O})$, one obtains

$$T_n\left(\Theta\left(\varphi_k\right)\right) = \Theta\left(\widetilde{T}_n\left(\varphi_k\right)\right) = \Theta\left(\tilde{t}_k(n)\varphi_k\right) = \tilde{t}_k(n)\Theta\left(\varphi_k\right)\;, \tag{6.29}$$

for all $n \in \mathbb{N}$. Hence Θ maps a Hecke basis of $L_0^2(X_\mathcal{O})$ to that part of a Hecke basis of $\mathcal{C}_d$ that spans $\operatorname{ran}\Theta \subset \mathcal{C}_d$. The Hecke eigenvalues are not changed under this map. In the same manner $\widetilde{\Theta}$ maps a Hecke basis $\{g_k;\; k \in \mathbb{N}\}$ of $\mathcal{C}_d$ onto a Hecke basis of $L_0^2(X_\mathcal{O})$,

$$\widetilde{T}_n\left(\widetilde{\Theta}\left(g_k\right)\right) = \widetilde{\Theta}\left(T_n g_k\right) = \widetilde{\Theta}\left(t_k(n)g_k\right) = t_k(n)\widetilde{\Theta}\left(g_k\right)\;, \tag{6.30}$$

again for all $n \in \mathbb{N}$. The theta-lifts Θ and $\widetilde{\Theta}$ therefore not only preserve Laplace eigenspaces and eigenvalues, but also Hecke eigenspaces and eigenvalues.

7 Newforms

In the previous section we saw that the theta-lift Θ maps a Hecke basis $\{\varphi_k;\ k \in \mathbb{N}\}$ of $L^2_0(X_{\mathcal{O}})$ into a Hecke basis $\{g_k;\ k \in \mathbb{N}\}$ of $\mathcal{C}_d$. Due to Corollary 5.1, a generic choice of the reference point z_0 ensures that $\operatorname{ran}\Theta \subset \mathcal{C}_d$ is contained in the linear span of the Hecke eigenforms g_k with non-vanishing first Fourier coefficients $c_k(1)$. Now a way to analyse such eigenforms is provided by the oldform-newform formalism as developed in [1] for holomorphic modular forms. However, the concepts and principal results carry over to Maaß cusp forms, see for example [9, §8.5].

Now let $a, m \in \mathbb{N}$, $m < d$, be such that $am|d$, and take some $h \in \mathcal{C}_m$. The inclusion $\Gamma_0(d) \subset \Gamma_0(m)$ implies that $\mathcal{C}_m \subset \mathcal{C}_d$ so that $h \in \mathcal{C}_d$, but also $h^{(a)} \in \mathcal{C}_d$ with $h^{(a)}(\tau) := h(a\tau)$. The linear span of all such forms $h^{(a)} \in \mathcal{C}_d$ that derive from all possible a, m is called the oldspace $\mathcal{C}_d^{old}$. Its orthogonal complement within $\mathcal{C}_d$ is the newspace $\mathcal{C}_d^{new}$, so that $\mathcal{C}_d = \mathcal{C}_d^{old} \oplus \mathcal{C}_d^{new}$. The oldspace is closed under the action of the Hecke operators T_n with $(n,d) = 1$. Since the latter are self-adjoint, the same is true for the newspace. One can therefore introduce a Hecke basis of $\mathcal{C}_d$ such that one part of this basis spans the oldspace, and the remaining part spans the newspace. A Hecke eigenform in the newspace is then called a newform. If h is a newform in $\mathcal{C}_m$, then $h^{(a)}$ is called an oldform in $\mathcal{C}_d$.

Let a, m be as above and consider a cusp form $h \in \mathcal{C}_m$ with Fourier expansion

$$h(\tau) = \sum_{n\neq 0} b(n)\, \sqrt{v}\, K_{ir}(2\pi|n|v)\, e^{2\pi i n u} \ . \tag{7.1}$$

Then $h^{(a)}$ has a Fourier expansion

$$\begin{aligned} h^{(a)}(\tau) &= \sum_{n\neq 0} b(n)\, \sqrt{av}\, K_{ir}(2\pi|na|v)\, e^{2\pi i n a u} \\ &= \sum_{m\equiv 0 \bmod a} c(m)\sqrt{v}\, K_{ir}(2\pi|m|v)\, e^{2\pi i m u} \ , \end{aligned} \tag{7.2}$$

with Fourier coefficients $c(m) = \sqrt{a}\, b(\frac{m}{a})$ if $m \equiv 0 \bmod a$, and $c(m) = 0$ otherwise. In particular, $a > 1$ implies that $c(1) = 0$. According to the non-holomorphic analogue of [1, Thm.5], $\mathcal{C}_d^{old}$ is spanned by oldforms, see also [9, p.129].

From the above considerations, and from the non-holomorphic analogue of [1, Lem.19] one concludes

Lemma 7.1 *Let $\{g_k;\ k \in \mathbb{N}\}$ be a Hecke basis of $\mathcal{C}_d$ and denote the Fourier coefficients of g_k by $c_k(n)$. Then $c_k(1) \neq 0$ iff $g_k \in \mathcal{C}_m^{new}$ for some divisor m of d.*

Since we know from Section 6, that the theta-lifts $\Theta(\varphi_k)$ of a Hecke basis $\{\varphi_k;\ k \in \mathbb{N}\}$ of $L^2_0(X_{\mathcal{O}})$ have non-vanishing first Fourier coefficients, we conclude that $\Theta(\varphi_k) \in \mathcal{C}_m^{new}$ for suitable $m|d$. An interesting question now arises: Are all theta-lifts newforms for $\Gamma_0(d)$, i.e. is it true that $\Theta(\varphi_k) \in \mathcal{C}_d^{new}$? In order to at least partially answer this question, we introduce the spectral counting functions for various classes of Hecke eigenforms,

$$\begin{aligned} N_d(\lambda) &:= \#\{\mu_k \leq \lambda\} \ , \\ N_d^{'}(\lambda) &:= \#\{\mu_k \leq \lambda :\ g_k \in \mathcal{C}_d^{new}\} \ , \\ N_d^{''}(\lambda) &:= \#\{\mu_k \leq \lambda :\ c_k(1) \neq 0\} \ , \\ N_d^{'''}(\lambda) &:= \#\{\Theta(\varphi_k) :\ \mu_k \leq \lambda\} \ . \end{aligned} \tag{7.3}$$

Due to Proposition 5.4 and (1.10), we know that

$$N_d^{'''}(\lambda) = N_{\mathcal{O}}(\lambda) \sim \frac{A_{\mathcal{O}}}{4\pi}\,\lambda =: \frac{\lambda}{12}\,\psi_{\mathcal{O}}\;, \quad \lambda \to \infty\;. \tag{7.4}$$

The asymptotics (1.15) together with (1.8) reads

$$N_d(\lambda) \sim \frac{\lambda}{12}\,\psi(d) \quad \lambda \to \infty\;, \tag{7.5}$$

where $\psi(d) := d\prod_{p|d}\left(1+\frac{1}{p}\right)$.

In order to determine the asymptotics of $N_d^{'}(\lambda)$ and $N_d^{''}(\lambda)$, respectively, we relate these to the known asymptotics (7.5) of $N_d(\lambda)$. For divisors m of d the oldspace $\mathcal{C}_m^{old}$ is built up by newforms of level $m^{'}|m$. This procedure can be traced back to the lowest possible level $m^{'} = 1$, i.e. to the full modular group. By definition, $\mathcal{C}_1 = \mathcal{C}_1^{new}$. Hence $\mathcal{C}_d$ is constructed from the newforms of all levels dividing d, including d itself. In this procedure one can form as many oldforms $h^{(a)}$ from $h \in \mathcal{C}_m^{new}$ as there are divisors a of $\frac{d}{m}$, and all these forms have the same Laplace eigenvalue. Denoting the number of positive divisors of $n \in \mathbb{N}$ by $\tau(n)$, one therefore obtains

$$N_d(\lambda) = \sum_{m|d} N_m^{'}(\lambda)\,\tau\left(\frac{d}{m}\right)\;, \tag{7.6}$$

which is the non-holomorphic analogue of [1, (6.6)]. According to [1, (6.7)] one can invert (7.6),

$$N_d^{'}(\lambda) = \sum_{m|d} N_m(\lambda)\,\beta\left(\frac{d}{m}\right)\;, \tag{7.7}$$

with $\beta(n) := \sum_{k|n}\mu(k)\mu(\frac{n}{k})$, and $\mu(n)$ is the Möbius function. Asymptotically for $\lambda \to \infty$, the r.h.s. of (7.7) yields

$$N_d^{'}(\lambda) \sim \frac{\lambda}{12}\,\psi^{'}(d) \quad \text{with} \quad \psi^{'}(d) = \sum_{m|d}\psi(m)\,\beta\left(\frac{d}{m}\right)\;. \tag{7.8}$$

The arithmetic function ψ is multiplicative in the sense that $\psi(nm) = \psi(n)\,\psi(m)$ if $(n,m) = 1$. We recall the following result from elementary number theory:

Lemma 7.2 *1. Let f and g be two arithmetic functions such that*

$$f(n) = \sum_{d|n} g(d)\;. \tag{7.9}$$

Then f is multiplicative iff g is multiplicative.

2. If $\tau(n)$ is the number of positive divisors of n, and f is an arithmetic function, then

$$g(n) := \sum_{d|n} f(d)\,\tau\left(\frac{n}{d}\right) = \sum_{d'|n}\sum_{d|d'} f(d). \tag{7.10}$$

In particular, by applying part 1. twice, we get that f is multiplicative iff g is.

An evaluation of both sides of (7.6) asymptotically results in

$$\psi(d) = \sum_{m|d}\psi^{'}(m)\,\tau\left(\frac{d}{m}\right)\;. \tag{7.11}$$

Part 2. of Lemma 7.2 then shows that $\psi^{'}$ is multiplicative since ψ is. We therefore only need the values of $\psi^{'}$ on prime powers. An easy calculation gives

$$\psi^{'}(p^e) = \begin{cases} p-1, & e=1, \\ p^2-p-1, & e=2, \\ p^{e-3}(p-1)^2(p+1), & e \geq 3, \end{cases} \tag{7.12}$$

A similar procedure will now be applied to the counting function $N_d^{''}(\lambda)$. As an immediate consequence of Lemma 7.1 we observe that

$$N_d^{''}(\lambda) = \sum_{m|d} N_m^{'}(\lambda) \ . \tag{7.13}$$

The asymptotic behaviour (7.8) of $N_d^{'}(\lambda)$ then yields for $\lambda \to \infty$

$$N_d^{''}(\lambda) \sim \frac{\lambda}{12} \psi^{''}(d) \quad \text{with} \quad \psi^{''}(d) = \sum_{m|d} \psi^{'}(m) \ . \tag{7.14}$$

Since $\psi^{'}$ is multiplicative, so is $\psi^{''}$ according to part 1. of Lemma 7.2. Therefore $\psi^{''}$ is determined by its values on prime powers,

$$\psi^{''}(p^e) = \begin{cases} p, & e=1 \ , \\ p^{e-2}(p-1)(p+1), & e \geq 2 \ . \end{cases} \tag{7.15}$$

In summary, we have

Proposition 7.1 *Let $N_d^{'}(\lambda)$ and $N_d^{''}(\lambda)$ be defined as above. Then for $\lambda \to \infty$*

$$N_d^{'}(\lambda) \sim \frac{\lambda}{12} \prod_{p|d} \psi^{'}(p^{e_p}) \ , \tag{7.16}$$

and

$$N_d^{''}(\lambda) \sim \frac{\lambda}{12} \prod_{p|d} \psi^{''}(p^{e_p}) \ , \tag{7.17}$$

when $\psi^{'}$ and $\psi^{''}$ are given by (7.12) *and* (7.15) *respectively, and $d = \prod p^{e_p}$ is the prime factorisation of d.*

Therefore, the asymptotic fraction of theta-lifts that are newforms in $\mathcal{C}_d$ is bounded from above by

$$F_{\mathcal{O}} := \lim_{\lambda \to \infty} \frac{N_d^{'}(\lambda)}{N_d^{'''}(\lambda)} = \psi_{\mathcal{O}}^{-1} \prod_{p|d} \psi^{'}(p^{e_p}) \ . \tag{7.18}$$

In case $\mathcal{O}$ is a maximal order the following consideration shows that $F_{\mathcal{O}} = 1$. It is therefore possible that in this case all theta-lifts are newforms. The discriminant $d = d(\mathcal{O})$ is given by the discriminant of the quaternion algebra, i.e. by the product of the ramified primes,

$$\prod_{p|d} \psi^{'}(p^{e_p}) = \prod_{p|d} \psi^{'}(p) = \prod_{p|d} (p-1) \ . \tag{7.19}$$

On the other hand, [25, Cor. IV.1.8] yields

$$\psi_{\mathcal{O}} = \phi(d) = \prod_{p|d} (p-1) \ , \tag{7.20}$$

so that the r.h.s. of (7.18) is equal to 1.

We also remark that Proposition 7.1 in conjunction with Proposition 5.4 implies that

$$\prod_{p|d} \psi''(p^{e_p}) \geq \psi_{\mathcal{O}}, \tag{7.21}$$

since $c_1 \neq 0$ for theta-lifts by the proof of Corollary 5.1. This puts a restriction on the least possible level d for a fixed order $\mathcal{O}$. In many cases one can conclude that $d = d(\mathcal{O})$ is best possible, in particular this is true for all maximal orders.

Now let $\mathcal{O}$ be a non-maximal order. Hence there exists an order $\mathcal{O}'$ in A containing $\mathcal{O}$ as a proper suborder. In this case $d' = d(\mathcal{O}')$ is a proper divisor of $d = d(\mathcal{O})$. Since $L_0^2(X_{\mathcal{O}'}) \subset L_0^2(X_{\mathcal{O}})$ one can lift $\varphi \in L_0^2(X_{\mathcal{O}'})$ either to $\Theta'(\varphi) \in \mathcal{C}_{d'}$, or to $\Theta(\varphi) \in \mathcal{C}_d$.

Proposition 7.2 *Let $\mathcal{O}_2 \subseteq \mathcal{O}_1$ be two Eichler orders in an indefinite quaternion algebra over $\mathbb{Q}$, and let $\varphi \in L_0^2(X_{\mathcal{O}_1}) \subseteq L_0^2(X_{\mathcal{O}_2})$. Fix a base point $z_0 \in \mathcal{H}$ and let Θ_1 and Θ_2 be the theta-lifts of $L_0^2(X_{\mathcal{O}_1})$ and $L_0^2(X_{\mathcal{O}_2})$, respectively (each corresponding to the same z_0). Then $\Theta_1(\varphi) = \Theta_2(\varphi)$.*

Proof. Let φ_k be a function in a Hecke basis of $L_0^2(X_{\mathcal{O}_1})$. We may assume that $\varphi_k(z_0) \neq 0$. Take an arbitrary $m \in \mathbb{N}$, such that $(m, d(\mathcal{O}_2)) = 1$. Since $d(\mathcal{O}_1)|d(\mathcal{O}_2)$, this also implies $(m, d(\mathcal{O}_1)) = 1$. Then the action of the Hecke operators $\widetilde{T}_m$ (T_m) is the same on $L_0^2(X_{\mathcal{O}_1})$ and $L_0^2(X_{\mathcal{O}_2})$ ($\mathcal{C}_{d(\mathcal{O}_1)}$ and $\mathcal{C}_{d(\mathcal{O}_2)}$). Hence by (6.29), we obtain

$$T_m(\Theta_1(\varphi_k)) = \tilde{t}_k(m)\Theta_1(\varphi_k) \text{ and } T_m(\Theta_2(\varphi_k)) = \tilde{t}_k(m)\Theta_2(\varphi_k). \tag{7.22}$$

Therefore, $\Theta_1(\varphi_k)$ and $\Theta_2(\varphi_k)$ have the same eigenvalues for all m with $(m, d(\mathcal{O}_2)) = 1$. By the non-holomorphic analogue of [1, Th.5], this implies that $\Theta_1(\varphi_k)$ and $\Theta_2(\varphi_k)$ are in the same class in the sense of [1]. Since the first Fourier coefficients $c_1 = 4\varphi_k(z_0) \neq 0$ agree, this implies that $\Theta_1(\varphi_k) = \Theta_2(\varphi_k)$. The result then follows by the linearity of Θ_1 and Θ_2. □

In the case of Proposition 7.2, we now denote the theta-lift of $L_0^2(X_{\mathcal{O}_1}) \cap L_0^2(X_{\mathcal{O}_2})$ simply as Θ.

Corollary 7.1 *If $\varphi \in L_0^2(X_{\mathcal{O}_1}) \cap L_0^2(X_{\mathcal{O}_2})$ for Eichler orders $\mathcal{O}_2 \subsetneq \mathcal{O}_1$, then $\Theta(\varphi)$ is not a newform in $\mathcal{C}_{d(\mathcal{O}_2)}$. In particular, if $\mathcal{O}$ is a non-maximal Eichler order, then there are $\varphi \in L_0^2(X_{\mathcal{O}})$ such that $\Theta(\varphi)$ is not a newform in $\mathcal{C}_{d(\mathcal{O})}$.*

Proof. This is an immediate consequence of Proposition 7.2, since Proposition 7.2 implies that $\Theta(\varphi) = \Theta_2(\varphi) \in \mathcal{C}_{d(\mathcal{O}_1)}$. □

In the case of maximal orders, we will use the fact that g is a newform iff it is an eigenform of all Hecke operators and all adjoints of Hecke operators, see Corollary 1 to Theorem 2 in [15]. This fact and the fact that the Hecke operators on $X_{\mathcal{O}}$ all are self-adjoint gives the following partial result:

Proposition 7.3 *Let V be the subspace of $L_0^2(X_{\mathcal{O}})$ spanned by Hecke forms φ such that the eigenvalue of the Laplacian of φ is simple. If φ is a Hecke form in V, then $\Theta(\varphi)$ is a newform.*

Proof. By (6.29), we get that $\Theta(\varphi)$ is an eigenform of all Hecke operators. Hence it is enough to show that $\Theta(\varphi)$ is an eigenform of all adjoints of Hecke operators.

Suppose that φ and ψ are Hecke forms in V. Then

$$\widetilde{\Theta}\Theta(\varphi) = c_\varphi \varphi \text{ and } \widetilde{\Theta}\Theta(\psi) = c_\psi \psi, \tag{7.23}$$

for some constants c_φ and c_ψ, since $\widetilde{\Theta}\Theta$ maps an eigenform to an eigenform with the same eigenvalue. From this, Lemma 5.2, Proposition 6.1 and the fact that all Hecke operators $\widetilde{T_p}$ are self-adjoint, we get

$$\begin{aligned}
\langle T_p\Theta(\varphi), \Theta(\psi)\rangle_{X_d} &= \left\langle \Theta(\widetilde{T_p}\varphi), \Theta(\psi)\right\rangle_{X_d} = \left\langle \widetilde{T_p}\varphi, \widetilde{\Theta}\Theta(\psi)\right\rangle_{X_\mathcal{O}} \\
&= c_\psi \left\langle \widetilde{T_p}\varphi, \psi\right\rangle_{X_\mathcal{O}} = c_\psi \left\langle \varphi, \widetilde{T_p}\psi\right\rangle_{X_\mathcal{O}} \\
&= \overline{c_\varphi}^{-1} c_\psi \left\langle \widetilde{\Theta}\Theta(\varphi), \widetilde{T_p}\psi\right\rangle_{X_\mathcal{O}} = \overline{c_\varphi}^{-1} c_\psi \left\langle \Theta(\varphi), \Theta(\widetilde{T_p}\psi)\right\rangle_{X_d} \\
&= \overline{c_\varphi}^{-1} c_\psi \langle \Theta(\varphi), T_p\Theta(\psi)\rangle_{X_d}.
\end{aligned} \tag{7.24}$$

Since $T_p^* = T_p$ for p such that $(p, d) = 1$, we get that $c_\varphi = c_\psi \in \mathbb{R}$. From this we derive $T_p^* = T_p$ on $\Theta(V)$. Hence all Hecke forms in $\Theta(V)$ are also eigenforms of T_p^* and therefore newforms. $\square$

Corollary 7.2 *Let V be as in Proposition 7.3. Then $\widetilde{\Theta}\Theta$ is a real multiple of the identity on V.*

References

[1] A. O. L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[2] J. Brzezinski. Arithmetical quadratic surfaces of genus 0, 1. *Math. Scand.*, 46:183–208, 1980.

[3] J. Brzezinski. A characterization of Gorenstein orders in quaternion algebras. *Math. Scand.*, 50:19–24, 1982.

[4] J. Brzezinski. On orders in quaternion algebras. *Comm. Algebra*, 11:501–522, 1983.

[5] I. Gelfand, M. Graev, and I. Pyatetskii-Shapiro. *Representation Theory and Automorphic Functions.* Saunders, Philadelphia, 1969.

[6] D. A. Hejhal. *The Selberg Trace Formula for $PSL(2,\mathbb{R})$ vol. 1.* Lecture Notes in Mathematics 548. Springer-Verlag, Berlin-Heidelberg-New York, 1976.

[7] D. A. Hejhal. *The Selberg Trace Formula for $PSL(2,\mathbb{R})$ vol. 2.* Lecture Notes in Mathematics 1001. Springer-Verlag, Berlin-Heidelberg-New York, 1983.

[8] D. A. Hejhal. A classical approach to a well-known spectral correspondence on quaternion groups. In D. Chudnovsky, G. Chudnovsky, H. Cohn, and M. Nathanson, editors, *Number Theory, New York 1983-84*, Lecture Notes in Mathematics 1135. Springer-Verlag, Berlin-Heidelberg-New York, 1985.

[9] H. Iwaniec. *Introduction to the Spectral Theory of Automorphic Forms.* Biblioteca de la Revista Matemática Iberoamericana, Madrid, 1995.

[10] H. Jacquet and R. Langlands. *Automorphic Forms on GL(2).* Lecture Notes in Mathematics 114. Springer-Verlag, Berlin-Heidelberg-New York, 1970.

[11] S. Johansson. Genera of Riemann surfaces corresponding to quaternion orders. Research Report No. 1994-37, Dept. of Mathematics, Chalmers University of Technology and Göteborg University, 1994.

[12] Y. Kitaoka. *Arithmetic of Quadratic Forms.* Cambridge University Press, Cambridge, 1993.

[13] S. Lang. *Algebraic Number Theory.* 2nd edition, Springer-Verlag, Berlin-Heidelberg-New York, 1994.

[14] W. Luo, Z. Rudnick, and P. Sarnak. On Selberg's eigenvalue conjecture. *Geometric and Functional Analysis*, 5:387–401, 1995.

[15] T. Miyake. On automorphic forms on GL_2 and Hecke operators. *Ann. Math.*, 94:174–189, 1971.

[16] T. Miyake. *Modular Forms.* Springer-Verlag, Berlin-Heidelberg-New York, 1989.

[17] S. Niwa. Modular forms of half integral weight and the integral of certain theta-functions. *Nagoya Math. J.*, 56:147–161, 1974.

[18] K. Ribet. Sur les variétés abéliennes à multiplications réelles. *C. R. Acad. Sci. Paris*, 291:121–123, 1980.

[19] R. Schoen, S. Wolpert, and S. Yau. Geometric bounds on the low eigenvalues of a compact surface. In *Proc. Symp. in Pure Math. 36*, 279–285, Amer. Math. Soc., Providence, 1980.

[20] A. Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Symp. in Pure Math. 8*, 1–15, Amer. Math. Soc., Providence, 1965.

[21] H. Shimizu. Theta series and automorphic forms on GL(2). *J. Math. Soc. Japan*, 24:638–683, 1972.

[22] G. Shimura. On modular forms of half-integral weight. *Ann. Math.*, 97:440–481, 1973.

[23] T. Shintani. On construction of holomorphic cusp-forms of half-integral weight. *Nagoya Math. J.*, 58:83–126, 1975.

[24] C. L. Siegel. Indefinite quadratische Formen und Modulfunktionen. In *Studies and Essays Presented to R. Courant on his 60:th Birthday.* Interscience, 1948.

[25] M.-F. Vignéras. *Arithmétique des Algèbres de Quaternions.* Lecture Notes in Math. 800. Springer-Verlag, Berlin-Heidelberg-New York, 1980.