# AN APPLICATION OF AN EXPLICIT TRACE FORMULA TO A WELL–KNOWN SPECTRAL CORRESPONDENCE ON QUATERNION GROUPS

ANDREAS STRÖMBERGSSON

This text contains the detailed proofs of the results announced in the paper [S6]. We refer to that paper for a general introduction and further comments.

## Contents



## 1. Recollection of known facts, and statement of results

Let $B$ be a quaternion algebra over $\mathbb{Q}$. We use $N_B$ and $\mathrm{tr}_B$ to denote the reduced norm $B \to \mathbb{Q}$ and the reduced trace $B \to \mathbb{Q}$ respectively. Cf. [M2, p. 200], [W1, pp. 168-169]. We let $\overline{\beta}$ denote the conjugate of an element $\beta \in B$. Thus $\mathrm{tr}_B(\beta) = \beta + \overline{\beta}$ and $N_B(\beta) = \beta\overline{\beta} = \overline{\beta}\beta$.

We recall that a quaternion algebra over any field $F$ is either a division algebra, or else it is isomorphic to $M_2(F)$ [W1, IX, Thm. 1], [V, p. 6, Cor. 2.4].

*Date*: April, 2000.

We let $d_B$ be the *reduced* discriminant of $B$ [V, p. 58(bottom)], [M2, p. 211(bottom)][1]. Then $d_B$ is square-free; it is the product of all primes $p$ such that $B$ ramifies over $\mathbb{Q}_p$, i.e. such that $B_p = B \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is a division algebra. We recall that for each square free number $d \in \mathbb{Z}^+$, there is exactly one quaternion algebra $B$ over $\mathbb{Q}$ with $d_B = d$, up to isomorphisms. Also, $d_B > 1$ if and only if $B$ is a division algebra. Cf. [V, p.74, Thm. 3.1 and below].

We will assume that $B$ is indefinite, which means that $B_\infty := B \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic to $M_2(\mathbb{R})$. We will *fix,* once and for all, *one* such isomorphism. By [V, p. 74, Thm. 3.1], $B$ being indefinite means that the reduced discriminant $d_B$ has an *even* number of prime factors.

We recall the definition of an order:

**Definition 1.1.** *If $A$ is a finite dimensional algebra over $\mathbb{Q}$, with identity element $1_A$, then a subset $\mathcal{O} \subset A$ is called an* order in $A$ *if the following three conditions are satisfied:*

*(i). $\mathcal{O}$ is a finitely generated $\mathbb{Z}$-module;*

*(ii). $\mathcal{O}$ contains a basis of $A$ over $\mathbb{Q}$;*

*(iii). $\mathcal{O}$ is a subring of $A$, and $1_A \in \mathcal{O}$.*

We remark that in this situation, $\mathcal{O}$ has a $\mathbb{Z}$-basis consisting of $n$ $\mathbb{Q}$-linearly independent elements ($n = \dim_{\mathbb{Q}} A$). Also:

(1.1) If $A'$ is a subalgebra of $A$, then $A' \cap \mathcal{O}$ is an order in $A'$.

From now on, we let $R$ be an order in our quaternion algebra $B$. Concerning the reduced norm and trace, we then have:

$$N_B(\beta) \in \mathbb{Z} \quad \text{and} \quad \mathrm{tr}_B(\beta) \in \mathbb{Z} \quad \text{for all} \quad \beta \in R. \tag{1.2}$$

Cf., e.g., [H5, Prop. 3.1].

We use $d_R$ to refer to the (reduced) discriminant of the order $R$. Thus

$$d_R = \sqrt{|\det\left[\mathrm{tr}_B(\xi_j \xi_k)\right]|},$$

where $\xi_1, \xi_2, \xi_3, \xi_4$ is any $\mathbb{Z}$-basis of $R$. $d_R$ is always an integer, independent of the choice of $\xi_1, ..., \xi_4$, and divisible by $d_B$. We have $d_R = d_B$ if and only if $R$ is a *maximal* order in $B$ [E, pp. 53,49($q_2 = 1$),19], [V, pp. 84–85].

Now let $\Gamma_R = \{\beta \in R \mid N_B(\beta) = 1\}$. Through our fixed isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$, $\Gamma_R$ is identified with a subgroup of $\mathrm{SL}(2, \mathbb{R})$. By [M2, pp. 209(bottom)-210], $\Gamma_R$ is actually a Fuchsian group. $\Gamma_R$ is

[1]The $d_B$ in [M2] is the square of ours.

cocompact if and only if $B$ is a division algebra, i.e. if and only if $d_B > 1$. We will assume that this is the case.

The spectral correspondence that we will herein focus on is a correspondence between the eigenfunctions of the Laplace-Beltrami operator $D = y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ on $L_2(\Gamma_R \setminus \mathcal{H})$ and the discrete eigenfunctions of $D$ on $L_2(\Gamma_0(d_R) \setminus \mathcal{H})$.

In [H5], this correspondence is obtained as an integral transform $\Theta$, the kernel of which is constructed as a special case of the [famous] *Siegel's theta function.* Next, we give a quick review of this construction, following the presentation in [BJ1, pp. 52–53], except for some small changes of notation. (Cf. also Section 7).

We *fix* a $\mathbb{Z}$-basis $\xi_1, \xi_2, \xi_3, \xi_4$ of $R$. Then there is a unique $\mathsf{B} \in \mathrm{GL}_4(\mathbb{R})$ such that our already fixed isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$ is given by (for $k_1, k_2, k_3, k_4 \in \mathbb{R}$):

(1.3)
$$k_1\xi_1 + k_2\xi_2 + k_3\xi_3 + k_4\xi_4 \mapsto \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \text{where} \quad \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \mathsf{B}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}.$$

We then let

$$S := \mathsf{B}^t \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \mathsf{B}. \tag{1.4}$$

Also, for $L_1, L_2 \in \mathrm{SL}(2,\mathbb{R})$, define $A(L_1, L_2) \in M_4(\mathbb{R})$ through

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

For $z = x + iy \in \mathcal{H}$, define $M_z \in \mathrm{SL}(2,\mathbb{R})$ through

$$M_z := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix} = \begin{pmatrix} y^{\frac12} & xy^{-\frac12} \\ 0 & y^{-\frac12} \end{pmatrix}.$$

Then, for $w, z \in \mathcal{H}$, define $P_{zw} \in M_4(\mathbb{R})$ by

$$P_{zw} := \mathsf{B}^t \cdot \Big(A(M_z^{-1}, M_w^{-1})^t \cdot A(M_z^{-1}, M_w^{-1})\Big) \cdot \mathsf{B}. \tag{1.5}$$

We refer to [H5] and [BJ1] for motivating comments to these definitions. Here we merely remark that by construction, $S$ is a symmetric matrix that satisfy (for $k_1, k_2, k_3, k_4 \in \mathbb{R}$):

$$\begin{pmatrix} k_1 & k_2 & k_3 & k_4 \end{pmatrix} \cdot S \cdot \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = 2 \cdot N_B\big(k_1\xi_1 + k_2\xi_2 + k_3\xi_3 + k_4\xi_4\big).$$

Cf. [BJ1, (4.3)]. In particular, by (1.2), $S$ is "even" (i.e. $s_{jj} \in 2\mathbb{Z}$ and $s_{ij} \in \mathbb{Z}$ for $1 \leq i, j \leq 4$). Also, for any $z, z_0 \in \mathcal{H}$, $P_{zz_0}$ is a "majorant" for $S$, i.e. $P_{zz_0}$ is symmetric, positive definite, and $P_{zz_0} S^{-1} P_{zz_0} = S$.

Finally, for $z_0, z, \tau \in \mathcal{H}$, we let:

$$\theta(z;\tau) := (\operatorname{Im}\tau) \cdot \sum_{m \in \mathbb{Z}^4} \exp\Big[\pi i \cdot m^t \big[(\operatorname{Re}\tau) S + i(\operatorname{Im}\tau) P_{zz_0}\big] m\Big]. \tag{1.6}$$

We call the "hidden" variable $z_0$ the *reference point*; it is kept *fixed* below.

Cf. [H5, pp. 136-137, 157-158]. Notice that $\theta(z;\tau)$ differs from the theta function used in [H5] by the extra factor $\operatorname{Im}\tau$.

This function has the following invariance properties:

**Theorem 1.2.**

$$\theta(Tz;\tau) = \theta(z;\tau), \qquad \textit{for each } T \in \Gamma_R$$
$$\theta(z;T\tau) = \theta(z;\tau), \qquad \textit{for each } T \in \Gamma_0(d_R).$$

*Proof.* This is [BJ1, Prop. 4.1]. Cf. also [H5, Prop. 4.2, p. 158(a,b)]. □

Let $\mathcal{F}_{\Gamma_R} \subset \mathcal{H}$ be a fundamental polygon for $\Gamma_R \setminus \mathcal{H}$. From the above theorem it is clear that if $f : \mathcal{H} \to \mathbb{C}$ is any $\Gamma_R$−invariant function (continuous, say), then the function

$$\Theta(f)(\tau) := \int_{\mathcal{F}_{\Gamma_R}} \theta(z;\tau) f(z)\, d\mu(z)$$

will not depend on the choice of $\mathcal{F}_{\Gamma_R}$, and it will be $\Gamma_0(d_R)$−invariant.

This defines the integral transform $\Theta$. Cf. [H5, (4.9)], [BJ1, (5.1)].

In this situation, if $f$ is a nonconstant eigenfunction of the Laplace-Beltrami operator $D$, then $\Theta(f)$ is *also* an eigenfunction of $D$, with the *same eigenvalue*, and $\Theta(f)$ is a *cusp form* on $\Gamma_0(d_R)$. This is proved in [H5] (for the general case, cf. [BJ1, Prop. 5.1]) by giving an explicit formula for the Fourier expansion of $\Theta(f)$ at the cusp $\infty$. When $R$ is a maximal order in $B$, this formula can be expressed in a particularly nice way, using the *Hecke operators* on $\Gamma_R$.[2] We restate this formula below (in Theorem 1.3), after devoting some paragraphs to introducing the Hecke operators on $\Gamma_R$.

We first consider the general case of an arbitrary cofinite Fuchsian group $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$ and a subset $H \subset \mathrm{GL}^+(2,\mathbb{R})$ such that

$$(1.7) \qquad H\Gamma = \Gamma H = H$$

$$(1.8) \qquad [H:\Gamma] < \infty, \text{ i.e. } H = \bigsqcup_{i=1}^{d} \Gamma\alpha_i \text{ for some } \alpha_i \in H \quad (d < \infty)$$

In this case, we can define $\mathcal{M}$, a *modular correspondence*, as an operator acting on $f \in L_2(\Gamma \setminus \mathcal{H})$ or $f \in C(\Gamma \setminus \mathcal{H})$ by[3]

$$(1.9) \qquad (\mathcal{M}f) = \sum_{i=1}^{d} f(\alpha_i z).$$

One can then easily prove that $\mathcal{M}f$ does not depend on the choice of $\alpha_i$'s in (1.8), that $\mathcal{M}f$ is again $\Gamma-$invariant, and that $\mathcal{M}$ is a bounded linear operator on $L_2(\Gamma \setminus \mathcal{H})$. Cf. [S5, Lemma 2.1, Remark 8.5] (only trivial changes are required to treat $\mathrm{SL}(2,\mathbb{R})$, $\mathrm{GL}^+(2,\mathbb{R})$ instead of $\mathrm{PSL}(2,\mathbb{R})$, $\mathrm{PGL}^+(2,\mathbb{R})$). $\mathcal{M}$ is trivially seen to commute with the Laplace-Beltrami operator $D$ on $C^\infty(\Gamma \setminus \mathcal{H})$.

Throughout the rest of this section:

We let $R$ be a *maximal* order in $B$.

We recall that all maximal orders in $B$ are conjugate [M2, Thm. 5.2.12]. Also: $d_R = d_B$.

Now for our quaternion group $\Gamma_R$, and any $n \in \mathbb{Z}^+$, we can take

$$(1.10) \qquad H_n = \{\beta \in R \mid N_B(\beta) = n\} \subset \mathrm{GL}^+(2,\mathbb{R});$$

the last inclusion is again given through our fixed isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$. Notice that $H_n$ corresponds to "$T(n)$" in [M2, (5.3.12)].

[2]Notice that for non-maximal Eichler orders (level $N > 1$), [BJ1, §6] contains errors, since the definition of Hecke operators given there (in (6.3), (6.4)) is *not* the same as in [M2] when $(n,N) > 1$; see [M2, §5.3, esp. (5.3.4), (5.3.12); notice the restriction $a \in \mathbb{Z}_p^\times$ in (5.3.4)].

[3]Recall in (1.9) and (1.12) that the action of any $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2,\mathbb{R})$ on $z \in \mathbb{C}$ is defined simply by $\alpha z := (az+b)/(cz+d)$. Thus if $\det \alpha < 0$, as in (1.12), then $\alpha$ interchanges $\mathcal{H}$ and $\overline{\mathcal{H}}$.

Clearly $H_n\Gamma_R = \Gamma_R H_n = H_n$, and from [M2, §5.3], $[H_n : \Gamma_R] < \infty$, and we can define the operator $\widetilde{T}_n$ on $L_2(\Gamma_R \setminus \mathcal{H})$ by

$$(1.11)\quad \widetilde{T}_n := \frac{1}{\sqrt{n}} \cdot [\text{the modular correspondence coming from } H_n].$$

Notice that $\widetilde{T}_1$ is simply the identity operator.

We need one more operator. For $f \in L_2(\Gamma_R \setminus \mathcal{H})$ we define

$$(1.12)\qquad (\mathcal{P}f)(z) = f(\eta\overline{z}), \qquad \text{where } \eta \in R \text{ and } N_B(\eta) = -1.$$

Cf. [H5, (8.12)]. For the existence of such $\eta$, see [S2, Lemma 1.4] or [M2, Thm. 5.2.10 ($N = 1$), Lemma 5.2.3(3)]. One easily checks that $\mathcal{P}f$ does not depend on the choice of $\eta$, and is $\Gamma_R$−invariant. Also, $\mathcal{P}$ is a bounded linear operator of $L_2(\Gamma_R \setminus \mathcal{H})$ of norm 1 (use that $\eta(\overline{\mathcal{F}_{\Gamma_R}})$ is a fundamental domain for $\eta\Gamma_R\eta^{-1} = \Gamma_R$). As such,

$$(1.13)\qquad \mathcal{P} \text{ is self adjoint, and } \mathcal{P} \circ \mathcal{P} = \mathrm{Id}.$$

$$(1.14)\qquad \mathcal{P} \circ \widetilde{T}_n = \widetilde{T}_n \circ \mathcal{P} \quad \text{for each } n \in \mathbb{Z}^+.$$

(See [S5, pp. 8–9]. Use $\eta^2 \in \Gamma_R$ and $\eta H_n \eta^{-1} = H_n$)

At last, we can now state the afore-mentioned result from [H5]. For any cofinite Fuchsian group $\Gamma$, and $\lambda \geq 0$, we use $A(\Gamma \setminus \mathcal{H}, \lambda)$ to denote the finite dimensional space of discrete eigenfunctions with eigenvalue $\lambda$, i.e.

$$(1.15)\qquad A(\Gamma \setminus \mathcal{H}, \lambda) := \{\phi \in C^\infty \cap L_2(\Gamma \setminus \mathcal{H}) \mid D\phi + \lambda\phi \equiv 0\}.$$

Notice that if $\Gamma$ is any congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, and $\lambda > 0$, then $A(\Gamma\backslash\mathcal{H}, \lambda)$ can only contain cusp forms (use [O1] or [H6][4] in conjunction with [H4, pp. 327(note 15), 78(claim 9.6), 284(line 1), 328(line -2)]).

**Theorem 1.3.** *If* $\varphi \in A(\Gamma_R \setminus \mathcal{H}, \lambda)$, $\lambda = r^2 + \frac{1}{4} > 0$, *then* $\Theta(\varphi) \in A(\Gamma_0(d_B)\backslash\mathcal{H}, \lambda)$, *i.e.* $\Theta(\varphi)$ *is a cusp form on* $\Gamma_0(d_B)$ *with the eigenvalue* $\lambda$. *The Fourier expansion of* $\Theta(\varphi)$ *at the cusp at infinity is given by*

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

*where*

$$c(n) = 4 \cdot \begin{cases} \left(\widetilde{T}_n\varphi\right)(z_0) & \text{if } n \geq 1 \\ \left(\mathcal{P}\widetilde{T}_{|n|}\varphi\right)(z_0) & \text{if } n \leq -1 \end{cases}.$$

[4] In this paper we will mainly be concerned with Hecke congruence subgroups $\Gamma_0(N)$ of square free level $N$, and then [H4, p. 535(Lemma 4.6)] can be used instead of [O1] or [H6].

*Proof.* This is [BJ1, Prop. 5.1], for $R$ =a maximal order. Cf. also [H5, (8.9)]. ☐

The purpose of the present paper is to study the *range* and the question of *injectivity* of the correspondence $\Theta : A(\Gamma_R \setminus \mathcal{H}, \lambda) \to A(\Gamma_0(d_R) \setminus \mathcal{H}, \lambda)$. We will always keep our assumption that $R$ is a maximal order, i.e. $d_R = d_B$. As suggested in [H5, p. 186, item 2] and developed further in [BJ1], the answer to the range question will involve the so called *newforms* for the congruence group $\Gamma_0(d_B)$. We now recall the definitions related to this concept. More details will be given in Section 4.

Taking $\lambda \geq 0, N \in \mathbb{Z}^+$, we introduce the notation

$$A(N, \lambda) := A(\Gamma_0(N) \setminus \mathcal{H}, \lambda). \tag{1.16}$$

For $\lambda > 0$, this is the space of cusp forms on $\Gamma_0(N)$ with eigenvalue $\lambda$, by the remark following (1.15). Clearly, if $M \mid N$, then $A(M, \lambda) \subset A(N, \lambda)$. For $l \in \mathbb{Z}^+$ we define the operator $\delta_l$, acting on functions $f : \mathcal{H} \to \mathbb{C}$ invariant under some congruence group, by

$$(\delta_l f)(z) = \frac{1}{\sqrt{l}} \cdot f(lz). \tag{1.17}$$

Then $\delta_l$ is a linear map from $A(N, \lambda)$ into $A(lN, \lambda)$ (cf. Lemma 4.4(e) below, clearly it is true also for $\lambda = 0$). We now define:

**Definition 1.4.** $A_{old}(N, \lambda)$, *the* oldspace *of level $N$ and eigenvalue $\lambda$, is the linear span of the spaces in*

$$\{\delta_l A(M, \lambda) \mid \quad M, l \in \mathbb{Z}^+, \quad lM \mid N, \quad M \neq N\}.$$

Cf. [M2, p. 162(lines 13–17)]. Notice that this is a subspace of $A(N, \lambda)$.

**Definition 1.5.** $A_{new}(N, \lambda)$, *the* newspace *of level $N$ and eigenvalue $\lambda$, is the orthogonal complement (in the usual inner product) of $A_{old}(N, \lambda)$ in $A(N, \lambda)$.*

We will then prove:

**Theorem 1.6.** *For every eigenfunction $\varphi \in A(\Gamma_R \setminus \mathcal{H}, \lambda)$ with eigenvalue $\lambda > 0$, $\Theta(\varphi) \in A_{new}(d_B, \lambda)$. Also, if the reference point $z_0 \in \mathcal{H}$ is chosen outside a certain set of (Lebesgue) measure zero, then for each $\lambda > 0$, $\Theta$ is a bijection between $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ and $A_{new}(d_B, \lambda)$.*

The proof of this is based on a comparison of [classical] trace formulas. This approach was suggested in [H5]. We will evaluate the "geometric side" in the trace formulas for $\Gamma_R$ and for the newspaces of $\Gamma_0(d_B)$, and show that these two are equal. Actually, we will need to use *trace formulas for Hecke operators*; using the "ordinary" trace formula we can show that

$$\dim A(\Gamma_R \setminus \mathcal{H}, \lambda) = \dim A_{\text{new}}(d_B, \lambda) \qquad \text{for each } \lambda > 0, \tag{1.18}$$

but it doesn't seem to help us to settle the questions about injectivity and surjectivity of the map $\Theta$.

Part of the program to use trace formulas for Hecke operators to prove bijectivity of $\Theta$ was carried through in [BJ2]. In this paper, (1.18) was demonstrated, and Theorem 1.6 was proved for each $\lambda > 0$ for which $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ (or $A_{\text{new}}(d_B, \lambda)$) is *one-dimensional*. But without this assumption, neither injectivity nor surjectivity could be proved, and even the question whether the range of $\Theta$ indeed belongs to $A_{\text{new}}(d_B, \lambda)$ was left unsettled. These questions are resolved in the present paper by considering Hecke operators not only of prime order, but of *composite* order, and by utilizing the *multiplicativity* of the Hecke eigenvalues. (The details of this argument are given in Section 6.)

Before closing this section, we will formulate a more precise version of the main theorem, stating how $\Theta$ maps a *Hecke basis* in $L_2(\Gamma_R \setminus \mathcal{H})$. To explain this concept, we need a lemma:

**Lemma 1.7.** *The Hecke operators $\widetilde{T}_n$ ($n \in \mathbb{Z}^+$) satisfy:*

*(a)* $\widetilde{T}_n^* = \widetilde{T}_n$ *on* $L_2(\Gamma_R \setminus \mathcal{H})$.

*(b)* $\widetilde{T}_n \circ \widetilde{T}_m = \widetilde{T}_m \circ \widetilde{T}_n$ *for all* $n, m$.

*(c)* $\widetilde{T}_n \circ \widetilde{T}_m = \widetilde{T}_{nm}$ if $(n, m) = 1$.

*(d)* $\widetilde{T}_p \circ \widetilde{T}_{p^e} = \begin{cases} \widetilde{T}_{p^{e+1}} & \text{if } p \mid d_B \\ \widetilde{T}_{p^{e+1}} + \widetilde{T}_{p^{e-1}} & \text{if } p \nmid d_B \end{cases}$ *for $p$ prime, $e \geq 1$.*

*Proof.* We have $H_n^{-1} = n^{-1} \cdot H_n$ in $\mathrm{GL}^+(2, \mathbb{R})$, since $\beta^{-1} = n^{-1}\overline{\beta}$ for $\beta \in H_n$, and $\overline{R} = R$ (follows from (1.2)). Now our assertion (a) follows from [S5, (2.5), Lemma 2.2]. Recall the definition of a Hecke algebra in [M2, §2.7]. Notice that multiplication in the Hecke algebra corresponds to composition[5] of the corresponding modular correspondences. In [M2, (5.3.12)] (see also [M2, (5.3.4),(5.3.5); $N = 1$]), the element "$T(n)$" in the Hecke algebra "$\mathcal{R}(\Gamma_R, \Delta_R)$" is the same as our

[5] *Reversed* composition, since our modular correspondences act from the left (1.9). But this makes no difference, since the Hecke algebra here is commutative.

$H_n$ (1.10). Now our assertions (b),(c),(d) follow from (1.11) and [M2, Cor. 5.3.7, (5.3.13);$T(p,p)$]. $\square$

Now since $\mathcal{P}$ and all $\widetilde{T}_n$ commute with the Laplace-Beltrami operator $D$, they preserve each eigenspace $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ ($\lambda \geq 0$), and on each such space, their restrictions form a commutative family of self adjoint operators, by (1.13), (1.14), Lemma 1.7(a,b). Hence there is a simultaneous diagonalization: For any $\lambda \geq 0$ such that $A(\Gamma_R \setminus \mathcal{H}, \lambda) \neq \{0\}$ there is an orthonormal basis $\{\widetilde{f}_1, ..., \widetilde{f}_D\}$ of $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ such that

$$\mathcal{P}\widetilde{f}_j = \widetilde{\epsilon}_j \cdot \widetilde{f}_j, \qquad \widetilde{T}_n \widetilde{f}_j = \widetilde{\tau}_j(n) \cdot \widetilde{f}_j \tag{1.19}$$

for each $j \in \{1, ..., D\}$, $n \in \mathbb{Z}^+$. From (1.13) and Lemma 1.7(a), we then get $\widetilde{\epsilon}_j \in \{1, -1\}$ and $\widetilde{\tau}_j(n) \in \mathbb{R}$.

We will call such a basis $\{\widetilde{f}_1, ..., \widetilde{f}_D\}$ a *Hecke basis* in $A(\Gamma_R \setminus \mathcal{H}, \lambda)$.

In (2.8), (2.9) we will define the Hecke operators $T_n$ ($n \in \mathbb{Z}^+$) on $L_2(\Gamma_0(d_B) \setminus \mathcal{H})$, and in (4.8) we define the "mirror operator" $V$ simply by $[Vf](z) = f(-\overline{z})$.

**Theorem 1.8.** *Take $\lambda > 0$ such that $A(\Gamma_R \setminus \mathcal{H}, \lambda) \neq \{0\}$ (or equivalently, $A_{new}(d_B, \lambda) \neq \{0\}$). Let $\{\widetilde{f}_1, ..., \widetilde{f}_D\}$ be a Hecke basis in $A(\Gamma_R \setminus \mathcal{H}, \lambda)$, with eigenvalues as in* (1.19)*. We then have*

$$V\big[\Theta(\widetilde{f}_j)\big] = \widetilde{\epsilon}_j \Theta(\widetilde{f}_j) \quad \textit{and} \quad T_n\big[\Theta(\widetilde{f}_j)\big] = \widetilde{\tau}_j(n) \cdot \Theta(\widetilde{f}_j)$$

*for all $j \in \{1, ..., D\}$, $n \in \mathbb{Z}^+$. Also, $\Theta : A(\Gamma_R \setminus \mathcal{H}, \lambda) \to A_{new}(d_B, \lambda)$ is a bijection if and only if the reference point $z_0$ satisfies $\widetilde{f}_j(z_0) \neq 0$ for each $j \in \{1, ..., D\}$. When this is the case, $\{\Theta(\widetilde{f}_1), ..., \Theta(\widetilde{f}_D)\}$ is an orthogonal basis in $A_{new}(d_B, \lambda)$, and each $\Theta(\widetilde{f}_j)$ is a nonzero multiple of a newform (see Definition 4.5 below).*

Finally, we mention a consequence of Theorem 1.8, a certain "multiplicity one" result:

**Theorem 1.9.** *Take $L \in \mathbb{Z}^+$. Assume that $g_1, g_2$ are nonzero $L_2(\Gamma_R \setminus \mathcal{H})$-functions such that*

$$\widetilde{T}_p g_1 = \widetilde{\tau}_p \cdot g_1, \qquad \widetilde{T}_p g_2 = \widetilde{\tau}_p \cdot g_2, \qquad (\widetilde{\tau}_p \in \mathbb{R})$$

*for each prime $p$ not dividing $L$. Then $g_2$ is a multiple of $g_1$, and both these functions belong to $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ for some $\lambda \geq 0$, and they are eigenfunctions to* all *Hecke operators $\widetilde{T}_n$, and to $\mathcal{P}$.*

## 2. Explicit trace formula for Hecke operators

We will start by recalling the Selberg trace formula for a modular correspondence on a general cofinite Fuchsian group $\Gamma$. In the case where $\Gamma$ is cocompact, this formula was computed in [H3, Ch. 5, Thm. 3.3]. The case where $\Gamma$ has $\kappa \geq 1$ cusps was treated in [S5].

We again consider the setting on p. 5 of a subset $H \subset \mathrm{GL}^+(2,\mathbb{R})$ satisfying (1.7) and (1.8), and the corresponding modular correspondence $\mathcal{M}$, a bounded operator on $L_2(\Gamma \setminus \mathcal{H})$ given by (1.9).

We will always assume $-1 \in \Gamma$ (and thus $H = -H$).

We will use $A(\lambda)$ as a short-hand for $A(\Gamma \setminus \mathcal{H}, \lambda)$, as defined in (1.15). Since $\mathcal{M}$ commutes with the Laplace-Beltrami operator $D$, it preserves each eigenspace $A(\lambda)$. Hence $\mathrm{Tr}\ \mathcal{M}_{|A(\lambda)}$ is well defined. Of course, $\mathrm{Tr}\ \mathcal{M}_{|A(\lambda)} = 0$ if $A(\lambda) = \{0\}$. As usual, we define $r_\lambda$ for $\lambda \geq 0$ by $\lambda = \frac{1}{4} + r_\lambda^2, \quad r_\lambda \in i[-\frac{1}{2}, 0] \cup [0,\infty)$.

Take an arbitrary test function $h(r)$ subject to the following hypotheses:

$$\begin{cases} \text{(a) } h(r) \text{ is analytic on } |\mathrm{Im}\, r| \leq \frac{1}{2} + \delta \text{ for some } \delta > 0; \\ \text{(b) } h(-r) = h(r); \\ \text{(c) } |h(r)| \leq A\,(1 + |\mathrm{Re}\, r|)^{-2-\delta} \qquad \text{in the strip from (a).} \end{cases} \tag{2.1}$$

Then, the *trace formula for the modular correspondence* $\mathcal{M}$ can be expressed as follows [S5, (8.1), Thm. 8.3, Remark 8.5][6]:

(2.2)

$$\begin{aligned} &\sum_{\lambda \geq 0} h(r_\lambda) \cdot \mathrm{Tr}\ \mathcal{M}_{|A(\lambda)} \\ &= \left| \left\{ \delta > 0 \mid \begin{pmatrix} \delta & 0 \\ 0 & \delta \end{pmatrix} \in H \right\} \right| \cdot \frac{\mu(\Gamma \setminus \mathcal{H})}{4\pi} \cdot \int_{-\infty}^{\infty} r h(r) \tanh(\pi r)\, dr \\ &+ \frac{1}{2} \sum_{\substack{T \in H^{-1}//\Gamma, \\ \text{hyperbolic,} \\ \text{doesn't fix cusp}}} \frac{\ln \mathcal{N}(T_0)}{\mathcal{N}(T)^{1/2} - \mathcal{N}(T)^{-1/2}} \cdot g(\ln \mathcal{N}(T)) \end{aligned}$$

[6] It is easy to translate from the slightly different setting $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$, $H \subset \mathrm{PGL}^+(2,\mathbb{R})$ that was used in [S5]. One uses the assumption $-1 \in \Gamma$, and [H3, p. 377(Lemma 4.2)].

$$+\frac{1}{2}\sum_{\substack{T\in H^{-1}//\Gamma,\\ \text{elliptic}}}\frac{1}{2M_T\cdot\sin\theta(T)}\cdot\int_{-\infty}^{\infty}\frac{e^{-2\theta(T)r}}{1+e^{-2\pi r}}\cdot h(r)\,dr$$

$$+\sum_{i,j}^{(p)}\left[(C(r_{ij})-\ln 2)\cdot g(0)+\frac{1}{4}h(0)-\frac{1}{2\pi}\int_{-\infty}^{\infty}h(r)\frac{\Gamma'(1+ir)}{\Gamma(1+ir)}\,dr\right]$$

$$+\frac{1}{2}\sum_{i,j}^{(h)}\frac{\sqrt{\gamma_{ij}}}{|x_{ij}-y_{ij}|}\cdot\left(\sum_{\substack{v\\ \text{mod }|x_{ij}-y_{ij}|}}\ln\mathbf{c}\left(N_j\Gamma N_j^{-1},\frac{v+y_{ij}r_{ij}}{y_{ij}-x_{ij}}\right)\right)\cdot g(\ln\gamma_{ij})$$

$$+\frac{1}{2}\sum_{i,j}^{(h)}\sqrt{\gamma_{ij}}\cdot\int_{|\ln\gamma_{ij}|}^{\infty}g(u)\cdot\frac{e^{u/2}+e^{-u/2}}{e^{u/2}-e^{-u/2}+|\gamma_{ij}^{1/2}-\gamma_{ij}^{-1/2}|}\,du$$

$$+\frac{1}{4\pi}\int_{-\infty}^{\infty}h(r)\cdot\sum_{j,l,k}\xi_{lj}(\frac{1}{2}+ir)\varphi'_{jk}(\frac{1}{2}+ir)\overline{\varphi_{lk}(\frac{1}{2}+ir)}\,dr$$

$$-\frac{1}{4}h(0)\cdot\sum_{j,l}\varphi_{jl}(\frac{1}{2})\xi_{lj}(\frac{1}{2}).$$

If $\Gamma$ is cocompact, the last five lines are to be omitted. Cf. [H3, pp. 487(Thm. 3.3), 351–352(Thm. 5.1), 449(remark 9.4)] and [S5, Thm. 3.6].

We will now give a detailed explanation of the notation used in this formula.

First of all,

$$g(u)=\frac{1}{2\pi}\int_{-\infty}^{\infty}h(r)e^{-iru}\,dr,\qquad u\in\mathbb{R},$$

$\mu(\Gamma\setminus\mathcal{H})$ denotes the hyperbolic area of some fundamental domain for $\Gamma$, and $H^{-1}//\Gamma$ is taken to be a complete set of representatives of the $\Gamma$-conjugacy classes in $H^{-1}$.

For $T\in\mathrm{GL}^+(2,\mathbb{R})$ hyperbolic, $\mathcal{N}(T)$ denotes the (unique) number $\gamma>1$ such that $T$, as a Möbius transformation, is conjugate to $z\mapsto\gamma z$. For hyperbolic $T\in H^{-1}$, not fixing a cusp, the centralizer $Z_\Gamma(T)=\{\sigma\in\Gamma\mid\sigma T=T\sigma\}$ is nontrivial, by [S5, Thm. 3.6]. We then choose $T_0\in\Gamma$ so that $Z_\Gamma(T)=\{\pm T_0^n\mid n\in\mathbb{Z}\}$.

For $T\in\mathrm{GL}^+(2,\mathbb{R})$ elliptic, $\theta(T)$ denotes the unique angle $\theta\in(0,\pi)$ such that $T$, as a Möbius transformation, is conjugate to $z\mapsto\dfrac{(\cos\theta)z-\sin\theta}{(\sin\theta)z+\cos\theta}$. For elliptic $T\in\mathrm{GL}^+(2,\mathbb{R})$, we also let $M_T=\frac{1}{2}\left|Z_\Gamma(T)\right|$.

We take $\{\eta_1, ..., \eta_\kappa\}$ to be a full system of inequivalent cusps of $\Gamma$, and choose $N_j \in \mathrm{SL}(2,\mathbb{R})$ such that $N_j\eta_j = \infty$ and $\Gamma_{\eta_j} = \{\pm N_j^{-1} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} N_j \mid n \in \mathbb{Z}\}$. See [S5, Remark 8.4].

We *fix* the choice of $\alpha_i$:s in (1.8). Then for each $i \in \{1, ..., d\}$, $j \in \{1, ..., \kappa\}$ we have
(2.3)

$$(\det \alpha_i)^{-1/2}\alpha_i = A_{ij} N_{k(i,j)}^{-1} \begin{pmatrix} 1 & r_{ij} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{ij}^{1/2} & 0 \\ 0 & \gamma_{ij}^{-1/2} \end{pmatrix} N_j, \qquad A_{ij} \in \Gamma$$

for some uniquely determined $k(i,j) \in \{1, ..., \kappa\}$ and $\gamma_{ij} > 0$, and for some $r_{ij} \in \mathbb{R}$ which is uniquely determined modulo $\mathbb{Z}$. In fact, $\gamma_{ij} \in \mathbb{Q}^+$, and we define $x_{ij}, y_{ij} \in \mathbb{Z}^+$ through $\gamma_{ij} = x_{ij}/y_{ij}$, $(x_{ij}, y_{ij}) = 1$. Cf. [S5, Lemma 1.11, (2.10), Remark 8.5].

We use $\sum_{i,j}^{(p)}$ and $\sum_{i,j}^{(h)}$ to denote summation over

$$\sum_{i,j}^{(p)} : \quad \{\langle i,j\rangle \mid i \in \{1, ..., d\},\, j \in \{1, ..., \kappa\},\, k(i,j) = j,\, \gamma_{ij} = 1\} \quad \text{and}$$

$$\sum_{i,j}^{(h)} : \quad \{\langle i,j\rangle \mid i \in \{1, ..., d\},\, j \in \{1, ..., \kappa\},\, k(i,j) = j,\, \gamma_{ij} \neq 1\},$$

respectively.

For any Fuchsian group $\Gamma'$ with parabolic fixpoint $p \in \mathbb{R}$ we denote

$$\mathbf{c}(\Gamma', p) = \inf\{c > 0 \mid \begin{pmatrix} 1 + cp & -cp^2 \\ c & 1 - cp \end{pmatrix} \in \Gamma'\}. \tag{2.4}$$

For $0 < v \leq 1$ we define $C_v$, a generalized Euler constant, by

$$\sum_{0 \leq n \leq x - v} \frac{1}{v + n} = \ln x + C_v + O_v\left(\frac{1}{\sqrt{x}}\right) \qquad \text{for } x > 0. \tag{2.5}$$

In particular, $C = C_1$ is the usual Euler constant. For $r \in \mathbb{R}$ we then define $C(r)$ by

$$C(r) = \begin{cases} \frac{1}{2}(C_v + C_{1-v}) - C & \text{if } r \notin \mathbb{Z},\ v \equiv -r \pmod 1,\ 0 < v < 1 \\ 0 & \text{if } r \in \mathbb{Z} \end{cases}.$$

We also define:

$$\xi_{lj}(s) = \sum_{\substack{i=1 \\ k(i,j)=l}}^{d} \gamma_{ij}^s. \tag{2.6}$$

Finally, $\varphi_{jk}(s)$ $(j,k \in \{1,...,\kappa\})$ denotes the elements of the scattering matrix, cf. [H4, p. 281], [S5, (2.9)].

We will now turn to the case of a Hecke operator on a Hecke congruence group. We *fix* some $N \in \mathbb{Z}^+$, and take $\Gamma = \Gamma_0(N)$. Notice that the eigenspaces $A(\lambda)$ now are the spaces $A(N,\lambda)$ in (1.16). We then define, for $n \in \mathbb{Z}^+$,

(2.7)
$$H_n = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{Z}, \quad ad-bc=n, \quad N \mid c, \quad (a,N)=1\},$$

It is clear that $\Gamma H_n = H_n\Gamma = H_n$. One checks by a straightforward computation that

$$H_n = \bigsqcup_{\substack{ad=n \\ (a,N)=1 \\ a,d\in\mathbb{Z}^+}} \bigsqcup_{b \bmod d} \Gamma \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \qquad \text{(disjoint union)}. \tag{2.8}$$

Cf. [M2, (4.5.1),(4.5.4),(4.5.14),(4.5.25)]. Hence $H_n$ satisfies the conditions (1.7), (1.8), and we may *define* the *Hecke operator* $T_n$ on $L_2(\Gamma_0(N)\backslash\mathcal{H})$ by

$$T_n = \frac{1}{\sqrt{n}} \cdot [\text{the modular correspondence coming from } H_n]. \tag{2.9}$$

This is familiar from [G2, pp. 58–59] and [M2, §4.5]. The factor $\frac{1}{\sqrt{n}}$ is customary when working with nonholomorphic cusp forms, cf. [I, p. 127]. Notice that $T_1$ is the identity operator.

We now *specialize* to the following case:

$$N \text{ square free, and } (n,N)=1. \tag{2.10}$$

In this case, we will investigate further the trace formula for the operator $T_n$.

Notice, that under the assumption (2.10), $(a,N)=1$ is always true as soon as $ad-bc=n$ and $N \mid c$ hold, and we can thus remove the constraint $(a,N)=1$ in (2.7) and (2.8).

Concerning the group $\Gamma=\Gamma_0(N)$, it is well known that $\mu(\Gamma\setminus\mathcal{H}) = \frac{\pi}{3}\cdot N\prod_{p|N}(1+\frac{1}{p})$. Cf. [M2, Thm. 4.2.5(2)]. Also, under (2.10), the number of cusps is $\kappa = 2^r = d(N)$, where $r$ are the number of primes dividing $N$, and $d$ is the usual divisor function (see (5.1)). We now give an explicit list of these cusps.

Let $A \in \mathbb{Z}^+$ be a divisor of $N$. Take $B = N/A$. Using the Chineese remainder theorem and [M2, Thm. 4.2.1(1)] (cf. the proof of Lemma 7.9 below), we can then select a matrix $W_A \in \mathrm{SL}(2,\mathbb{Z})$ such that

$$W_A \equiv \begin{cases} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \text{mod } A \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{mod } B \, , \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{mod } n \end{cases} \tag{2.11}$$

We take $W_A$ to be fixed once and for all. Then define

$$\eta_A = W_A(\infty), \qquad N_A = \begin{pmatrix} A^{-1/2} & 0 \\ 0 & A^{1/2} \end{pmatrix} W_A^{-1}. \tag{2.12}$$

**Lemma 2.1.** $\{\eta_A \mid A|N\}$ *contains exactly one representative from each cuspclass of* $\Gamma_0(N)$. *Also, the fixator subgroup* $\Gamma_{\eta_A}$ *equals* $\{\pm N_A^{-1} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} N_A \mid m \in \mathbb{Z}\}$

*Proof.* See [H4, p. 534, lemma 4.2]. (For this lemma, we only need the congruences modulo $A$ and $B$ in (2.11), not modulo $n$.) □

Thus we can use the divisors $A$ of $N$ as a substitute for the cusp-index $j \in \{1, ..., \kappa\}$ from the general case. Then our $N_A$:s are exactly the $N_j$:s from the general case.

We next give the decomposition (2.3):

**Lemma 2.2.** *For* $\langle a, d, b\rangle$ *as in the disjoint union* (2.8) *and* $A \mid N$ *we have in* (2.3)*:*

$$k = k(\langle a,d,b\rangle, A) = A, \qquad \gamma = \gamma(\langle a,d,b\rangle, A) = \frac{a}{d}$$
$$x = x(\langle a,d,b\rangle, A) = \frac{a}{(a,d)}, \qquad y = y(\langle a,d,b\rangle, A) = \frac{d}{(a,d)},$$
$$r = r(\langle a,d,b\rangle, A) \equiv \frac{bA^*}{d} \mod \mathbb{Z},$$

*where the notation is adapted to* (2.8) *in an obvious way, and where* $A^* \in \mathbb{Z}$ *is any multiplicative inverse of* $A$ *modulo* $d$.

*Proof.* Concerning the existence of $A^*$, notice that $(d, A) = 1$, since $d \mid n$, $A \mid N$, $(n, N) = 1$.

We have to prove that (2.3) holds for the above choice of $k, \gamma, r$, i.e. that

$$\frac{1}{\sqrt{n}} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} N_A^{-1} \begin{pmatrix} \sqrt{d/a} & 0 \\ 0 & \sqrt{a/d} \end{pmatrix} \begin{pmatrix} 1 & -r \\ 0 & 1 \end{pmatrix} N_A \in \Gamma_0(N).$$

But the product evaluates to

$$\frac{1}{n} \left[ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} W_A \begin{pmatrix} d & -Adr \\ 0 & a \end{pmatrix} W_A^{-1} \right],$$

and we only have to verify that the integer matrix within the brackets is actually

$$\equiv \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \pmod{n} \qquad \text{and} \qquad \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{nN}.$$

This is a direct consequence of $Adr \equiv b \pmod{d}$, $Adr \equiv 0 \pmod{A}$ and (2.11). □

With notation as above, notice that multiplication with $A^*$ simply permutes the residue classes modulo $d$. Using this, the $\sum^{(p)}$–sum in the trace formula (2.2) is evaluated just as in the case $N = 1$, see [S5, p. 51], except that we have to multiply with the number of cusps. We get 0 if $\sqrt{n} \notin \mathbb{Z}$, and otherwise:

$$\kappa \left[ \ln\left(\frac{\sqrt{n}}{2}\right) \cdot g(0) + \frac{1}{4} h(0) - \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r) \frac{\Gamma'(1+ir)}{\Gamma(1+ir)}\, dr \right].$$

To continue, we notice that $N_A \Gamma_0(N) N_A^{-1} = \Gamma_0(N)$. This is verified as a direct consequence of (2.11) and (2.12). Also, for each $p \in \mathbb{Q}^+$, $\mathbf{c}(\Gamma_0(N), p) = \inf\{c \in N\mathbb{Z}^+ \mid cp^2 \in \mathbb{Z}\}$, see (2.4).

Hence the first $\sum^{(h)}$–sum in (2.2) is (remember that we have to divide by $\sqrt{n}$, because of (2.9)):

$$\frac{1}{2\sqrt{n}} \sum_{A|N} \sum_{\substack{ad=n \\ a \neq d}} \sum_{b \bmod d} \frac{\sqrt{a/d} \cdot (a,d)}{|a-d|} g(\ln \frac{a}{d})$$

$$\times \sum_{v \bmod \frac{|a-d|}{(a,d)}} \ln \inf\{c \in N\mathbb{Z}^+ \mid c \left[ \frac{v(a,d) + dr}{d-a} \right]^2 \in \mathbb{Z}\}.$$

Here $dr \in \mathbb{Z}$, by the definition of $r$. The second line is, if we substitute $v$ for $v(a,d)$ :

$$\sum_{\substack{v \in (a,d)\mathbb{Z} \\ \bmod |a-d|\mathbb{Z}}} \ln \mathrm{LCM}\left(N, \left[\frac{d-a}{(d-a, v+dr)}\right]^2\right).$$

(LCM = least common multiple.)
Clearly, this sum only depends on the congruence class of $dr$ modulo $(a,d)$. When we add over $b$, $dr$ runs through a complete system of (integral) residues modulo $d$, and thus it runs through all the residues modulo $(a,d)$ exactly $d/(a,d)$ times. Hence our expression equals:

$$\begin{aligned}\frac{1}{2\sqrt{n}} \sum_{A|N} \sum_{\substack{ad=n \\ a\neq d}} & \frac{\sqrt{a/d}\cdot(a,d)}{|a-d|} g(\ln\frac{a}{d}) \\ & \times \frac{d}{(a,d)} \sum_{w \bmod |a-d|} \ln \mathrm{LCM}\left(N, \frac{(d-a)^2}{(d-a,w)^2}\right).\end{aligned}$$

Now use the formula $\mathrm{LCM}(x,y) = xy/(x,y)$ $(x, y \in \mathbb{Z}^+)$, and the fact that $N$ is square free. We get

$$d(N) \sum_{\substack{ad=n \\ a\neq d}} \left(\frac{1}{2}\ln N + \ln|a-d| - \frac{\ln X(|a-d|)}{|a-d|} - \frac{1}{2}\frac{\ln Y(|a-d|, N)}{|a-d|}\right) \cdot g(\ln\frac{a}{d}),$$

where we have defined

$$X(m) = \prod_{k \bmod m} (k,m), \qquad \text{and} \tag{2.13}$$

$$Y(m,N) = \prod_{k \bmod m} \left(N, \frac{m}{(k,m)}\right), \tag{2.14}$$

for $m \in \mathbb{Z}^+$.

We remark that

$$\begin{aligned}&\frac{\ln X(m)}{m} + \frac{1}{2}\frac{\ln Y(m,N)}{m} \\ &= \sum_{p|m} \left\{ \begin{array}{ll} 1 & \text{if } p \nmid N \\ \frac{p+1}{2} & \text{if } p \mid N \end{array} \right\} \left(\frac{1}{p} + \frac{1}{p^2} + \ldots + \frac{1}{(p^\infty, m)}\right) \cdot \ln p.\end{aligned} \tag{2.15}$$

Similarly, but easier, the second $\sum^{(h)}$–sum in (2.2) is seen to be

$$\frac{\kappa}{2}\sum_{\substack{ad=n\\ a\neq d}}\int_{|\ln(\frac{a}{d})|}^{\infty} g(u)\cdot\frac{e^{u/2}+e^{-u/2}}{e^{u/2}-e^{-u/2}+|\sqrt{a/d}-\sqrt{d/a}|}\,du$$

For the next to last term in (2.2) we first notice that by (2.6) and Lemma 2.2, $\xi_{A_1A_2}(s)\equiv 0$ when $A_1\neq A_2$, and

$$\xi_{AA}(s)=\sum_{ad=n}\sum_{b \bmod d}\left(\frac{a}{d}\right)^s=\sqrt{n}\cdot\sum_{ad=n}\left(\frac{a}{d}\right)^{s-1/2};$$

This means that for each $A\mid N$, the Eisenstein series $E_A(z,s)$ associated to the cusp $\eta_A$ is an eigenfunction to $T_n$, with eigenvalue $\sum_{ad=n}(a/d)^{s-1/2}$. Cf. [S5, (2.6), Lemma 2.6].

The sum in the next to last term in (2.2) now equals

$$\begin{aligned}&\sum_{A_1|N}\sum_{A_2|N}\xi_{A_1A_1}(\frac{1}{2}+ir)\varphi'_{A_1A_2}(\frac{1}{2}+ir)\overline{\varphi_{A_1A_2}(\frac{1}{2}+ir)}\\ &=\xi_{11}(\frac{1}{2}+ir)\cdot\mathrm{Tr}\left[\Phi'(\frac{1}{2}+ir)\overline{\Phi(\frac{1}{2}+ir)^t}\right]\end{aligned}$$

where $\Phi(s)=(\varphi_{A_1A_2})_{A_1,A_2|N}$ is the scattering matrix. Letting $\varphi(s)=\det\Phi(s)$, and using [S5, (7.10)] and [H4, p. 315(i)] we get:

$$\xi_{11}(\frac{1}{2}+ir)\cdot\frac{\varphi'(\frac{1}{2}+ir)}{\varphi(\frac{1}{2}+ir)}.$$

But $\varphi(s)$ is given in [H4, p. 538(4.7)] (notice [H4, p. 508(2.2)=(2.3)]). Using this, our expression evaluates to:

(2.16)

$$\begin{aligned}&\xi_{11}(\frac{1}{2}+ir)\cdot\kappa\left[2\ln\pi-\frac{\Gamma'(\frac{1}{2}+ir)}{\Gamma(\frac{1}{2}+ir)}-\frac{\Gamma'(\frac{1}{2}-ir)}{\Gamma(\frac{1}{2}-ir)}-2\frac{\zeta'(1+2ir)}{\zeta(1+2ir)}-2\frac{\zeta'(1-2ir)}{\zeta(1-2ir)}\right]\\ &-\xi_{11}(\frac{1}{2}+ir)\cdot\frac{\kappa}{2}\sum_{p|N}\left(\frac{2\ln p}{1-p^{-1-2ir}}+\frac{2\ln p}{1-p^{-1+2ir}}\right).\end{aligned}$$

To compute the next to last term in (2.2), this should be multiplied by $(4\pi\sqrt{n})^{-1}h(r)$, and then integrated over $r\in\mathbb{R}$. The contribution from the first line of (2.16) is evaluated exactly as in [S5, pp. 53-54], yielding

$$\kappa(\ln\pi)\sum_{ad=n} g(\ln(\frac{a}{d})) - \frac{\kappa}{2\pi}\int_{-\infty}^{\infty} h(r)\cdot\sum_{ad=n}\left(\frac{a}{d}\right)^{ir}\cdot\frac{\Gamma'(\frac{1}{2}+ir)}{\Gamma(\frac{1}{2}+ir)}\,dr$$
$$-\frac{\kappa}{2}\left(\sum_{ad=n}1\right)\cdot h(0) + 2\kappa\cdot\sum_{m=2}^{\infty}\sum_{ad=n}\frac{\Lambda(m)}{m}\cdot g\left(\ln(\frac{a}{d}) - 2\ln m\right)$$

(since $\Lambda(1)=0$).

The contribution from the second line of (2.16) is:

$$-\frac{1}{4\pi}\int_{-\infty}^{\infty} h(r)\sum_{ad=n}\left(\frac{a}{d}\right)^{ir}\kappa\sum_{p|N}\ln p\left(\frac{1}{1-p^{-1-2ir}}+\frac{1}{1-p^{-1+2ir}}\right)dr$$
$$\left\{h(r)\sum_{ad=n}(a/d)^{ir}\text{ is even}\right\}$$
$$= -\kappa\sum_{p|N}\ln p\sum_{ad=n}\frac{1}{2\pi}\int_{-\infty}^{\infty} h(r)\left(\frac{a}{d}\right)^{ir}\left(\frac{1}{1-p^{-1-2ir}}\right)dr$$
$$= -\kappa\sum_{p|N}\ln p\sum_{ad=n}\frac{1}{2\pi}\int_{-\infty}^{\infty} h(r)\left(\frac{a}{d}\right)^{ir}\left(1+\sum_{m=1}^{\infty}p^{-m}p^{-2mir}\right)dr$$
$$= -\kappa\ln N\cdot\sum_{ad=n} g\left(\ln(\frac{a}{d})\right) - \kappa\sum_{p|N}\sum_{m=1}^{\infty}\sum_{ad=n}\frac{\ln p}{p^m}\cdot g\left(\ln(\frac{a}{d}) - 2\ln p^m\right)$$

Finally, $\varphi_{AA}(\frac{1}{2}) = -1$ and $\xi_{AA}(\frac{1}{2}) = \sqrt{n}(\sum_{ad=n}1)$ for each $A \mid N$, hence the last term in (2.2) gives $(\kappa/4)h(0)(\sum_{ad=n}1)$. We also notice that the first term in the right hand side of (2.2) occurs if and only if $\sqrt{n}\in\mathbb{Z}$, and then its first factor is 1. Adding, we finally obtain the trace formula for the operator $T_n$ on $\Gamma=\Gamma_0(N)$ ($N, n\in\mathbb{Z}^+$, $(N,n)=1$, and $N$ square free):

$$
\begin{aligned}
&\sum_{\lambda\geq 0} h(r_\lambda)\cdot \mathrm{Tr}\ (T_n)_{|A(N,\lambda)} \\
&= \left[\text{if } \sqrt{n}\in\mathbb{Z}: \quad \frac{N\cdot\prod_{p|N}(1+\frac{1}{p})}{12\sqrt{n}}\int_{-\infty}^{\infty} r h(r)\tanh(\pi r)\,dr\right] \\
&+ \frac{1}{2\sqrt{n}}\cdot \sum_{\substack{T\in H_n^{-1}//\Gamma_0(N) \\ \text{hyperbolic,} \\ \text{doesn't fix cusp}}} \frac{\ln\mathcal{N}(T_0)}{\mathcal{N}(T)^{1/2}-\mathcal{N}(T)^{-1/2}}\cdot g(\ln\mathcal{N}(T)) \\
&+ \frac{1}{2\sqrt{n}}\cdot \sum_{\substack{T\in H_n^{-1}//\Gamma_0(N), \\ \text{elliptic}}} \frac{1}{2M_T\cdot\sin\theta(T)}\cdot\int_{-\infty}^{\infty}\frac{e^{-2\theta(T)r}}{1+e^{-2\pi r}}\cdot h(r)\,dr \\
&+ \left[\text{if } \sqrt{n}\in\mathbb{Z}: \quad d(N)\left[\ln\left(\frac{\pi\sqrt{n}}{2N}\right)\cdot g(0)+\frac{1}{4}h(0)-\frac{1}{2\pi}\int_{-\infty}^{\infty}h(r)\frac{\Gamma'(1+ir)}{\Gamma(1+ir)}\,dr\right]\right] \\
&+ d(N)\sum_{\substack{ad=n \\ a\neq d}}\Bigg(-\frac{1}{2}\ln N+\ln\pi+\ln|a-d|-\frac{\ln X(|a-d|)}{|a-d|} \\
&\qquad\qquad\qquad\qquad -\frac{1}{2}\frac{\ln Y(|a-d|,N)}{|a-d|}\Bigg)\cdot g(\ln\frac{a}{d}) \\
&+ \frac{d(N)}{2}\sum_{\substack{ad=n \\ a\neq d}}\int_{|\ln(\frac{a}{d})|}^{\infty} g(u)\cdot\frac{e^{u/2}+e^{-u/2}}{e^{u/2}-e^{-u/2}+|\sqrt{a/d}-\sqrt{d/a}|}\,du \\
&- \frac{d(N)}{2\pi}\int_{-\infty}^{\infty}h(r)\cdot\sum_{ad=n}\left(\frac{a}{d}\right)^{ir}\cdot\frac{\Gamma'(\frac{1}{2}+ir)}{\Gamma(\frac{1}{2}+ir)}\,dr-\frac{d(N)}{4}\left(\sum_{ad=n}1\right)\cdot h(0) \\
&+ d(N)\sum_{m=2}^{\infty}\begin{Bmatrix} 1 & \text{if } \exists p \mid N: m=p^r \\ 2 & \text{otherwise}\end{Bmatrix}\frac{\Lambda(m)}{m}\cdot\sum_{ad=n} g\left(\ln(\frac{a}{d})-2\ln m\right)
\end{aligned}
$$

Here $a,d\in\mathbb{Z}^+$. For $X(m)$ and $Y(m,N)$, see (2.13), (2.14) and (2.15). Recall that $d(N)=\sum_{e|N}1=\prod_{p|N}2$ is the number of cusps of $\Gamma_0(N)$.

This formula is a generalization of both [H4, p. 538 (4.6)] and [H2, (11.10)]; it reduces to the first when $n=1$, and to the second when $N=1$ and $n$ is a prime. Also, one can show that the recent computations in [CL] lead, after further evaluation and one minor correction, to the same result as ours[7].

The corresponding formula [BJ2, Thm. 4.5] (which is for the case $n = p$, a prime) contains certain errors[8], due to an inadequate treatment of the hyperbolic elements that fix cusps.

## 3. Counting the elliptic and hyperbolic conjugacy classes

In this section we show how to get more explicit expressions for the hyperbolic and elliptic sums in the trace formula for a Hecke operator on $\Gamma_0(N)$ or on a quaternion unit group $\Gamma_R$. For these cases, it is well known that the hyperbolic and elliptic conjugacy classes can be enumerated in terms of class numbers and fundamental units of quadratic field extensions of $\mathbb{Q}$. See eg. [M2, §6.5-6.7], [V, Ch. III, §5C]. We have chosen to base our exposition on [M2, §6.5-6.7]. We will use many facts about quaternion algebras from [M2, §5.2].

The key tool in the computation will be [M2, Lemma 6.5.2], which allows us to pass from the global counting problem to an adelic one. We apply it in (3.24). The ratio between the adelic and the global expression is here a certain class number for a quadratic field extension of $\mathbb{Q}$. Before we start, we will devote a few pages to recalling some basic facts and definitions related to such extensions.

The nonisomorphic quadratic extensions of $\mathbb{Q}$ are given by $\mathbb{Q}(\sqrt{d})$, where $d$ runs through all *fundamental discriminants*, i.e. all nonzero integers such that

$$\begin{aligned}(3.1)\quad &\Big[d \equiv 1 \pmod 4 \quad \wedge \quad d \text{ square free} \quad \wedge \quad d \neq 1\Big] \\ &\text{or } \Big[d \equiv 0 \pmod 4 \quad \wedge \quad \frac{d}{4} \not\equiv 1 \pmod 4 \quad \wedge \quad \frac{d}{4} \text{ square free}\Big].\end{aligned}$$

The norm and the trace of $\mathbb{Q}(\sqrt{d})$ over $\mathbb{Q}$ are given by

$$N(x + y\sqrt{d}) = x^2 - dy^2; \qquad \operatorname{tr}(x + y\sqrt{d}) = 2x, \qquad \text{for } x, y \in \mathbb{Q}.$$

---

[7]Compare [CL, (3.12)] to the last six lines of our formula. The last "+" in line 2 of (3.12) should be corrected to "−".

[8]A term $2^{\omega(m)}\hat{h}(\log p)\left[\log m - (p-1)^{-1}\log Y(p-1,m)\right]$ must be added in the right hand side of their formula.

Below, the quadratic field will arise as the splitting field over $\mathbb{Q}$ of a polynomial

$$P(X) = X^2 - tX + n \tag{3.2}$$

where $t, n \in \mathbb{Z}$. Notice that $P(X)$ is irreducible if and only if $\sqrt{t^2 - 4n} \notin \mathbb{Q}$. When this holds, there are unique numbers $l, d$ such that

$$t^2 - 4n = l^2 d, \qquad l \in \mathbb{Z}^+, \qquad d \text{ as in (3.1)}. \tag{3.3}$$

Then the splitting field of $P(X)$ is $\mathbb{Q}(\sqrt{d})$, and the zeros of $P(X)$ are

$$x_1 = \frac{t}{2} + l \cdot \frac{\sqrt{d}}{2}, \quad x_2 = \frac{t}{2} - l \cdot \frac{\sqrt{d}}{2}.$$

Given any $d$ as in (3.1) we let, for each $f \in \mathbb{Z}^+$ :

$$\mathfrak{r}[f] := \mathbb{Z} + f\omega\mathbb{Z}, \qquad \text{where } \omega = \frac{d + \sqrt{d}}{2}, \tag{3.4}$$

Then $\mathfrak{r}[1]$, $\mathfrak{r}[2]$, $\mathfrak{r}[3], \ldots$ are all the distinct orders in $\mathbb{Q}(\sqrt{d})$, and in particular, $\mathfrak{r}[1]$ is the unique maximal order [D, p. 146–147]. Since $\big[\mathfrak{r}[1] : \mathfrak{r}[f]\big] = f$, we may call $\mathfrak{r}[f]$ *the order of index $f$ in* $\mathbb{Q}(\sqrt{d})$. For later reference, we note that

(3.5) $\mathfrak{r}[f_1] \subset \mathfrak{r}[f_2]$ if.f. $f_2 \mid f_1$.

(3.6) For $P(X)$ as in (3.2) and $d, l$ as in (3.3), the roots $x_1, x_2$ of $P(X)$ satisfy $\mathbb{Z} + \mathbb{Z}x_1 = \mathbb{Z} + \mathbb{Z}x_2 = \mathfrak{r}[l]$.

For any order $\mathfrak{r}$ in $\mathbb{Q}(\sqrt{d})$, we let $\mathfrak{r}^1 := \{\alpha \in \mathfrak{r} \mid N(\alpha) = 1\}$. This is a group under multiplication. If $d$ is negative, $\mathfrak{r}^1$ has finite cardinality;

$$\big|\mathfrak{r}[f]^1\big| = \frac{\big|\mathfrak{r}[1]^1\big|}{\big[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\big]} = \begin{cases} 4 & \text{if } d = -4,\ f = 1 \\ 6 & \text{if } d = -3,\ f = 1 \\ 2 & \text{otherwise} \end{cases}. \tag{3.7}$$

Cf. [M2, p. 259(top)].

If $d$ is positive, then all groups $\mathfrak{r}[f]^1$ are infinite, and of finite index in $\mathfrak{r}[1]^1$. We define the proper fundamental unit[9] $\epsilon_d$ in $\mathbb{Q}(\sqrt{d})$ by $\epsilon_d = (x + y\sqrt{d})/2$, where $\langle x, y\rangle$ is the positive integer solution to $x^2 - dy^2 = 4$ for which $y$ is minimal. This $\epsilon_d$ always exists, $\epsilon_d \in \mathfrak{r}[1]^1$, and the groups $\mathfrak{r}[f]^1$ are given by

$$\mathfrak{r}[f]^1 = \left\{\pm \left(\epsilon_d^A\right)^k \mid k \in \mathbb{Z}\right\}, \quad \text{where } A = \big[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\big]. \tag{3.8}$$

Cf. [L1, Vol I, Satz 111 + Vol III, pp. 183-185], and [C3, Ch. VI, §§4-5].

[9]This need not be the same as the fundamental unit, cf. [L1, Vol III, Satz 884].

At this point, we need to recall some general definitions. For any finite dimensional algebra $A$ over $\mathbb{Q}$, and for each place $v$ of $\mathbb{Q}$ (i.e. $v = \infty$ or $v =$a prime) we define $A_v := A \otimes_{\mathbb{Q}} \mathbb{Q}_v$. In the case $A = B$, this was done already in Section 1. $A_v$ is an algebra over $\mathbb{Q}_v$ of dimension $= \dim_{\mathbb{Q}} A$. For $v = p$ a prime, *orders* in $A_p$ are defined as in Definition 1.1, replacing $\mathbb{Q}$ and $\mathbb{Z}$ with $\mathbb{Q}_p$ and $\mathbb{Z}_p$. Equivalently, an order in $A_p$ is[10] an open compact subring in $A_p$ containing $1_{A_p}$. If $\mathcal{O}$ is an order in $A$, we get an order in $A_p$ through the definition:

$$\mathcal{O}_p := \big[\text{the closure of } \mathcal{O} \text{ in } A_p\big] = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p. \tag{3.9}$$

Also, recall the definition of the adele ring $A_{\mathbb{A}}$ [M2, pp. 197-198], [W1, p. 62]. As a set, this is

$$A_{\mathbb{A}} = \{(\alpha_v) \in \prod_v A_v \mid \alpha_p \in \mathcal{O}_p \text{ for almost all primes } p\},$$

where $\mathcal{O}$ is any order[11] in $A$.

We now look at the special case where our algebra $A$ is the field $K = \mathbb{Q}(\sqrt{d})$, $d$ as above. We have:

$$K_\infty \cong \begin{cases} \mathbb{R} \oplus \mathbb{R} & \text{if } d > 0 \\ \mathbb{C} & \text{if } d < 0 \end{cases}.$$

To describe $K_p$ for primes $p$, it is convenient to use the Kronecker quadratic reciprocity symbol, see [L1, Vol I, p. 51, Def. 20]. We will only use it in the form "$(d/p)$", where $d$ is as in (3.1), and $p$ is a prime number. For $p$ odd, this is just the usual thing, but for $p = 2$, the definition is

$$\left(\frac{d}{2}\right) = \begin{cases} 1 & \text{if } d \equiv 1 \pmod 8 \\ 0 & \text{if } d \equiv 0 \pmod 4 \\ -1 & \text{if } d \equiv 5 \pmod 8 \end{cases}$$

Recall the definition of *unramified* / *fully ramified* extensions of commutative $p$-fields [W1, p. 15, Def. 4]. We have:

$$K_p \text{ is } \begin{cases} \cong \mathbb{Q}_p \oplus \mathbb{Q}_p & \text{if } (d/p) = 1 \\ \text{a fully ramified extension of } \mathbb{Q}_p & \text{if } (d/p) = 0 \\ \text{an unramified extension of } \mathbb{Q}_p & \text{if } (d/p) = -1 \end{cases}. \tag{3.10}$$

Cf. [W2, §6.5]. A related fact is that if $\sqrt{t^2 - 4n} \notin \mathbb{Q}$ in (3.2),(3.3), the polynomial $P(X)$ is irreducible over $\mathbb{Q}_p$ if and only if $(d/p) \neq 1$.

[10]To prove the equivalence: Use [W1, Ch. II, Prop. 5] and the fact that the closure of $\mathbb{Z}$ in $A_p$ is $\mathbb{Z}_p$.

[11]actually, any $\mathbb{Q}$-lattice in $A$ will work here.

The distinct orders in $K_p$ are

$$\mathfrak{o}[p^n] := \mathbb{Z}_p + p^n\omega\mathbb{Z}_p, \qquad n = 0, 1, 2, ...$$

Here $\omega$ is the same as in (3.4), through the obvious embedding $K \subset K_p$. We have $\mathfrak{o}[1] \supset \mathfrak{o}[p] \supset \mathfrak{o}[p^2]$... and $\big[\mathfrak{o}[1] : \mathfrak{o}[p^n]\big] = p^n$. Also, combining definitions (3.4) and (3.9) we get:

$$(\mathfrak{r}[f])_p = \mathfrak{o}[p^n] \qquad \text{where } n = \mathrm{ord}_p f. \tag{3.11}$$

For any order $\mathfrak{r}$ in $K = \mathbb{Q}(\sqrt{d})$ we define (cf. [M2, p. 246]):

$$\begin{aligned}
&\mathfrak{r}_\infty := K_\infty \\
&\mathfrak{r}_A := \prod_v \mathfrak{r}_v \subset K_A \\
&\mathfrak{r}_{\infty+}^\times := \{x \in K_\infty^\times \mid N(x) > 0\}. \\
&\mathfrak{r}_{A+}^\times := \{(\alpha_v) \in \mathfrak{r}_A^\times \mid N(\alpha_\infty) > 0\} = \mathfrak{r}_{\infty+}^\times \times \prod_p \mathfrak{r}_p^\times \subset \mathfrak{r}_A^\times.
\end{aligned}$$

In the last two lines, $N$ denotes the extension of the norm $N : K \to \mathbb{Q}$ to the regular norm in the algebra $K_\infty$ over $\mathbb{Q}_\infty$, cf. [W1, p. 53]. In $\prod_v$, $v$ runs through all places of $\mathbb{Q}$, and in $\prod_p$, $p$ runs through all primes. Also, for any ring $R$ with unit (eg. $\mathfrak{r}_A$ or $\mathfrak{r}_p$ above), we write $R^\times$ for the multiplicative group of the invertible elements of $R$.

We then define the *class number of* $\mathfrak{r}$ through (cf. [M2, p. 246]):

$$\mathbf{h}(\mathfrak{r}) := \big[K_A^\times : (\mathfrak{r}_{A+}^\times \cdot K^\times)\big]. \tag{3.12}$$

We conclude by some remarks on how this number is related to some other class numbers defined in the literature. These facts will not be needed for the proofs of Theorems 1.6, 1.8, 1.9.

Definition (3.12) actually applies for any order in *any* finite field extension $K$ of $\mathbb{Q}$. $\mathbf{h}(\mathfrak{r})$ is always a finite number. One can prove – this requires some work – that $\mathbf{h}(\mathfrak{r})$ is the same as the number of ideal classes in $\mathfrak{r}$, as defined in [D, pp. 131,132,136: " $h'$ "][12]. For $K = \mathbb{Q}(\sqrt{d})$ as above, one can also show (cf. [C2, §5.2]) that $\mathbf{h}(\mathfrak{r}[f])$ is equal to the number of inequivalent primitive [and if $d < 0$ : positive] quadratic forms $ax^2 + bxy + cy^2$ with discriminant $b^2 - 4ac = df^2$. In other words: $\mathbf{h}(\mathfrak{r}[f]) = h(df^2)$, where the right hand side is as in [L1, Vol I, pp. 127–]. For $\mathfrak{r} = \mathfrak{r}[1]$, the maximal order, $\mathbf{h}(\mathfrak{r})$ is simply the narrow class number

[12] This is also the same as the number of equivalence classes of invertible ideals of $\mathfrak{r}$ (cf. [N2, Ch. I, §12], [C2, §4.9]), using *narrow* equivalence (cf. [C2, Def. 5.2.7(2)]). This is not an obvious fact, since not every invertible ideal of $\mathfrak{r}$ is an ideal in Dedekind's sense.

$\mathbf{h}_K$ of $K = \mathbb{Q}(\sqrt{d})$, as in [L1, vol. III, p. 187]. Extending the proof of [M2, Thm. 6.7.2(1)] to general $d$, we see that

$$\mathbf{h}(\mathfrak{r}[f]) = \frac{\mathbf{h}_K \cdot f \cdot \prod_{p|f} \left(1 - \left(\frac{d}{p}\right) p^{-1}\right)}{\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}. \tag{3.13}$$

We now turn to our investigation of the elliptic and hyperbolic sums in the trace formulas from last section. We will treat two cases:

[I]. The case of a Hecke operator $T_n$ on $\Gamma_0(N)$ ($N, n \in \mathbb{Z}^+$, $N$ square free, $(N, n) = 1$);

[II]. The case of an arbitrary Hecke operator $\widetilde{T}_n$ ($n \in \mathbb{Z}^+$) on a quaternion unit group $\Gamma_R$, for a *maximal* order $R$ in an indefinite division quaternion algebra $B$ over $\mathbb{Q}$. Cf. (1.10), (1.11).

We will refer to these two cases as case [I] and case [II].

For case [I], the trace formula is given on p. 19. For case [II] it is given by the formula (2.2), multiplied by $n^{-1/2}$, taking $H = H_n$ from (1.10), and omitting the five last lines.

To be able to treat the two cases simultaneously, we will take, *in case [I]*:

$$B = M_2(\mathbb{Q}) \text{ and } R = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \subset M_2(\mathbb{Z}) \mid \quad N \mid c \right\}.$$

Then $B$ is an indefinite quaternion algebra over $\mathbb{Q}$, and $R$ is an order in $B$. Notice that $R$ is *not* a maximal order if $N \geq 2$. Clearly, $\Gamma_R = \Gamma_0(N)$. Here of course, we have chosen the isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$ to be the canonical one, coming from $B = M_2(\mathbb{Q})$. Notice that the constraint $(a, N) = 1$ can be removed in (2.7), since $(n, N) = 1$.

Hence *both in case [I] and in case [II],* we have

$$H_n = \{\beta \in R \mid N_B(\beta) = n\} \subset \mathrm{GL}^+(2, \mathbb{R}).$$

Notice that $H_n^{-1} = n^{-1} H_n$. This is proved using $\beta^{-1} = n^{-1}\overline{\beta}$ for $\beta \in H_n$, together with the fact $\overline{R} = R$, which is a consequence of (1.2).

We can therefore add over $T \in H_n // \Gamma_R$ instead of $T \in H_n^{-1} // \Gamma_R$ in the hyperbolic and elliptic sums that we are considering. Hence we have to evaluate:

$$
\begin{aligned}
(3.14) \qquad & \frac{1}{2\sqrt{n}} \cdot \sum_{\substack{T \in H_n//\Gamma_R \\ \text{hyperbolic,} \\ \text{doesn't fix cusp}}} \frac{\ln \mathcal{N}(T_0)}{\mathcal{N}(T)^{1/2} - \mathcal{N}(T)^{-1/2}} \cdot g(\ln \mathcal{N}(T)) \\
& + \frac{1}{2\sqrt{n}} \cdot \sum_{\substack{T \in H_n//\Gamma_R \\ \text{elliptic}}} \frac{1}{2M_T \cdot \sin \theta(T)} \cdot \int_{-\infty}^{\infty} \frac{e^{-2\theta(T)r}}{1 + e^{-2\pi r}} \cdot h(r)\, dr.
\end{aligned}
$$

Recall that $\mathrm{tr}_B(H_n) \subset \mathbb{Z}$ (by (1.2)). Hence we have a disjoint union

$$
(3.15) \qquad H_n = \bigsqcup_{t \in \mathbb{Z}} H_{n,t}, \qquad \text{where } H_{n,t} := \{\beta \in H_n \mid \mathrm{tr}_B(\beta) = t\}.
$$

Notice that each $H_{n,t}$ is closed under conjugation by $\Gamma_R$-elements. In fact, if $t^2 - 4n \neq 0$, then $H_{n,t}$ contains only *finitely* many distinct $\Gamma_R$-conjugacy classes. This follows from the next lemma, together with [H3, pp. 486–487(Prop. 3.1,3.2)] and [S5, Thm. 3.6, Lemma 6.2, Remark 8.5(last paragraph)][13].

**Lemma 3.1.** *Take* $T \in GL^+(2, \mathbb{R})$, *and assume* $n = \det T$ *and* $t = tr\, T$. *Then for the Möbius transformation* $T : \mathcal{H} \to \mathcal{H}$, *we have:*

$$
\begin{aligned}
&\text{if } t^2 - 4n > 0 \textit{ then } T \textit{ is hyperbolic, and } \mathcal{N}(T) = \frac{\left(|t| + \sqrt{t^2 - 4n}\right)^2}{4n}; \\
&\text{if } t^2 - 4n = 0 \textit{ then } T \textit{ is parabolic or identity;} \\
&\text{if } t^2 - 4n < 0 \textit{ then } T \textit{ is elliptic, and } \cos \theta(T) = \pm \frac{t}{2\sqrt{n}}.
\end{aligned}
$$

*Proof.* We quickly reduce to the case $T \in \mathrm{SL}(2, \mathbb{R})$, by replacing $T$ with $n^{-1/2}T$. Then if $t^2 - 4 > 0$, $T$ is hyperbolic, and since $T$ as a Möbius transformation is conjugate to $z \mapsto \mathcal{N}(T) \cdot z$, we have $|t| = \mathcal{N}(T)^{1/2} + \mathcal{N}(T)^{-1/2}$. Solving for $\mathcal{N}(T)$, and remembering $\mathcal{N}(T) > 1$, we get $\mathcal{N}(T) = \frac{1}{4}(|t| + \sqrt{t^2 - 4})^2$. If $t^2 = 4$, then $T$ is parabolic or identity. If $t^2 - 4 < 0$, then $T$ is elliptic, and as a Möbius transformation, $T$ is conjugate to $z \mapsto \dfrac{(\cos \theta(T))z - \sin \theta(T)}{(\sin \theta(T))z + \cos \theta(T)}$, hence $t = \pm 2 \cos \theta(T)$. □

*Remark* 3.2. Concerning the expression for the multiplier in the hyperbolic case, notice that if $d \in \mathbb{Z}^+$ satisfies (3.1), $x, y \in \mathbb{Q}^+$, and if $t$ and

[13] Notice that $\mathcal{N}(T_0) \geq c(\Gamma_R) > 1$ for all primitive hyperbolic $T_0 \in \Gamma_R$; cf. [H3, p. 7, Prop. 2.5] and [H4, p. 475, Thm. 3.5].

$n$ denotes the trace and the norm of $x+y\sqrt{d}\in\mathbb{Q}(\sqrt{d})$, then

$$\frac{(|t|+\sqrt{t^2-4n})^2}{4n}=\frac{(x+y\sqrt{d})^2}{n}.$$

*Proof.* For any $x,y\in\mathbb{Q}$, if $t=\mathrm{tr}\ (x+y\sqrt{d})$, $n=N(x+y\sqrt{d})$ :

$$\frac{(|t|+\sqrt{t^2-4n})^2}{4n}=\frac{\left(2|x|+\sqrt{4x^2-4(x^2-dy^2)}\right)^2}{4n}=\frac{(|x|+|y|\sqrt{d})^2}{n}.$$

$\square$

**Lemma 3.3.** *Assume case [I]. Take $t\in\mathbb{Z}$, and $T\in H_{n,t}$. Assume that $T$ is hyperbolic, i.e. $t^2-4n>0$. Then both fixpoints of $T$ are cusps of $\Gamma_0(N)$ if $\sqrt{t^2-4n}\in\mathbb{Q}$, otherwise no fixpoint of $T$ is a cusp of $\Gamma_0(N)$.*

*Proof.* We recall that the cusps of $\Gamma_0(N)$ are exactly the points in $\{i\infty\}\cup\mathbb{Q}$, by [M2, Cor. 1.5.5, Th. 4.1.3(2)]. Assume $T=\begin{pmatrix}a&b\\c&d\end{pmatrix}$. If $c=0$, then

$$\sqrt{t^2-4n}=\sqrt{(a+d)^2-4ad}=|a-d|\in\mathbb{Q},$$

and the two fixpoints of $T$ are seen to be $i\infty$ and $b/(d-a)$, i.e. cusps of $\Gamma_0(N)$. If $c\neq 0$ then the fixpoints of $T$ are the two real solutions $x$ to

$$\begin{aligned}&\frac{ax+b}{cx+d}=x\iff x^2+2\frac{d-a}{2c}x=\frac{b}{c}\\&\iff\left(x+\frac{d-a}{2c}\right)^2=\frac{4bc+(d-a)^2}{4c^2}\iff x=\frac{a-d}{2c}\pm\frac{\sqrt{t^2-4n}}{2|c|},\end{aligned}$$

which are both rational if $\sqrt{t^2-4n}\in\mathbb{Q}$, otherwise both irrational. $\square$

**Lemma 3.4.** *Assume case [I]. Take $t\in\mathbb{Z}$ such that $\sqrt{t^2-4n}\notin\mathbb{Q}$, and define $d,l$ as in* (3.3)*. Then $H_{n,t}$ is nonempty if and only if:*

$$\left[p\mid l\ \text{or}\ \left(\frac{d}{p}\right)\neq -1\right],\qquad \text{for each prime } p \text{ dividing } N.$$

*Proof.* By definition, $H_{n,t}\neq\emptyset$ if and only if there are some $a,b,c\in\mathbb{Z}$ such that $\begin{pmatrix}a&b\\Nc&t-a\end{pmatrix}$ has norm $n$. Hence $H_{n,t}\neq\emptyset$ if and only if there is some $a\in\mathbb{Z}$ such that $a(t-a)-n\equiv 0\pmod N$. Remember that $N$

is square free. Hence our congruence equation is solvable if and only if it is solvable modulo $p$ for each prime $p$ dividing $N$. Now if $p \neq 2$ :

$$\begin{aligned} &a(t-a)-n \equiv 0 \pmod{p} \Longleftrightarrow (2a-t)^2 \equiv t^2-4n \pmod{p} \\ &\Longleftrightarrow (2a-t)^2 \equiv l^2 d \pmod{p}. \end{aligned}$$

There are solutions $a \in \mathbb{Z}$ to this equation if and only if $\Big[p \mid l$ or $(d/p) \neq -1\Big]$.

The case $p=2$ needs special treatment:

$$a(t-a)-n \equiv 0 \pmod{2} \Longleftrightarrow (2a-t)^2 \equiv l^2 d \pmod{8}.$$

If $2 \mid l$, then (by (3.3)) $2 \mid t$, and the congruence equation is solvable. If $2 \nmid l$, then $l^2 \equiv 1 \pmod{8}$ and (by (3.3)) $t \equiv d \pmod{2}$, and hence if $d \equiv 1 \pmod{8}$ or $d \equiv 0 \pmod{4}$ the equation is solvable, but not if $d \equiv 5 \pmod{8}$. $\square$

For each prime $p$ we get an order $R_p$ in $B_p$ through (3.9). We see that in case [I], $R_p = \begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ N\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix}$, and in case [II], $R_p$ is a maximal order in $B_p$ [M2, Lemma 5.2.4(1)].

**Lemma 3.5.** *Assume case [II]. Take $t \in \mathbb{Z}$ such that $t^2-4n \neq 0$. Then $H_{n,t}$ is nonempty if and only if $\sqrt{t^2-4n} \notin \mathbb{Q}$ and $\left(\dfrac{d}{p}\right) \neq 1$ for $d$ as in* (3.3) *and for all primes $p$ dividing $d_B$, the reduced discriminant of $B$.*

*Proof.* We first show, that the given condition is in fact necessary and sufficient for the existence of *any* $T \in B$ such that $N_B(T) = n$ and $\mathrm{tr}_B(T) = t$. Define $P(X)$ as in (3.2). This polynomial has two distinct roots, since $t^2-4n \neq 0$. Hence by [M2, Lemma 5.2.5(1)⇔(2) and p. 200(bottom)] there exists a $T \in B$ with $N_B(T) = n$ and $\mathrm{tr}_B(T) = t$ if and only if $B \otimes_{\mathbb{Q}} K \cong M_2(K)$, where $K$ is the splitting field of $P(X)$ over $\mathbb{Q}$. If $\sqrt{t^2-4n} \in \mathbb{Q}$, then $K = \mathbb{Q}$, and there are no $T$ of desired type, since $B \ncong M_2(\mathbb{Q})$ in case [II].

Now assume $\sqrt{t^2-4n} \notin \mathbb{Q}$, so that $P(X)$ is irreducible over $\mathbb{Q}$. For each place $v$ of $\mathbb{Q}$, let $K_{(v)}$ be the splitting field of $P(X)$ over $\mathbb{Q}_v$. Recall that $B_v = B \otimes_{\mathbb{Q}} \mathbb{Q}_v$. By [M2, Lemma 5.2.5(1)⇔(3)], an element $T \in B$ of desired type exists if and only if

$$B_v \otimes_{\mathbb{Q}_v} K_{(v)} \cong M_2(K_{(v)}) \tag{3.16}$$

for all places $v$ of $\mathbb{Q}$. When $v = \infty$, or when $v$ is a prime not dividing $d_B$, then $B_v \cong M_2(\mathbb{Q}_v)$, and (3.16) holds. Hence we only have to check (3.16) for $v$ running through the prime divisors $p$ of $d_B$. For these $p$ we have $B_p \ncong M_2(\mathbb{Q}_p)$, but if $K_{(p)} \ncong \mathbb{Q}_p$, then $B_p \otimes_{\mathbb{Q}_p} K_{(p)} \cong M_2(K_{(p)})$, by [W1, Ch. XII, Th. 2–Cor. 2] (cf. [V, Ch. II, Thm. 1.3]). Hence a $T \in B$ of desired type exists if and only if $P(X)$ is irreducible over $\mathbb{Q}_p$ for each prime $p$ dividing $d_B$, i.e. if and only if $(d/p) \neq 1$ for each such $p$.

Finally, if $T \in B$ is such that $N_B(T) = n$ and $\mathrm{tr}_B(T) = t$, we will show that $H_{n,t}$ is nonempty. We will first construct $\gamma = (\gamma_v) \in B_A^\times$ such that $\gamma_p^{-1} T \gamma_p \in R_p$ for all primes $p$. For all primes $p$ except finitely many, we have $T \in R_p$, and may thus take $\gamma_p = 1$. We also let $\gamma_\infty = 1 \in B_\infty^\times$. Now assume that $p$ is a prime such that $T \notin R_p$. Then by [M2, Lemma 5.2.3(1)], $B_p$ is not a division algebra, i.e. $B_p \cong M_2(\mathbb{Q}_p)$, and by [M2, Lemma 5.2.3(2)] we can assume that $R_p$ corresponds to $M_2(\mathbb{Z}_p)$ under this isomorphism. We have $T \notin \mathbb{Q}_p$, since $t^2 - 4n \neq 0$. Hence by [M2, Lemma 5.2.2(2) and p.200 (last line)] we can define $\gamma_p \in B_p^\times$ so that $\gamma_p^{-1} T \gamma_p = \begin{pmatrix} t & n \\ -1 & 0 \end{pmatrix} \in R_p$.

Next, by [M2, Thm. 5.2.11, Lemma 5.2.3(3)] we have $B_A^\times = B^\times \cdot \left(\mathrm{GL}^+(2,\mathbb{R}) \times \prod_p R_p^\times\right)$, and thus $\gamma = bu$ for some $b \in B^\times$, $u = (u_v) \in \mathrm{GL}^+(2,\mathbb{R}) \times \prod_p R_p^\times$. Then $b^{-1}Tb = u\gamma^{-1}T\gamma u^{-1}$, and for each prime $p$ we have $u_p \gamma_p^{-1} T \gamma_p u_p^{-1} \in u_p R_p u_p^{-1} = R_p$, since $u_p \in R_p^\times$. Hence $b^{-1}Tb \in R$ [M2, Lemma 5.2.4(4)]. We also have $N_B(b^{-1}Tb) = n$ and $\mathrm{tr}_B(b^{-1}Tb) = t$. Hence $b^{-1}Tb \in H_{n,t}$. $\square$

Using (3.15) and the last four lemmas, we can rewrite (3.14) as follows. First notice that in the hyperbolic case in Lemma 3.1, we get

$$\mathcal{N}(T)^{1/2} - \mathcal{N}(T)^{-1/2} = \frac{|t| + \sqrt{t^2 - 4n}}{2\sqrt{n}} - \frac{2\sqrt{n}}{|t| + \sqrt{t^2 - 4n}} = \frac{\sqrt{t^2 - 4n}}{\sqrt{n}}.$$

In the elliptic case in Lemma 3.1, notice that (since $\theta(T) \in (0, \pi)$)

$$\sin\theta(T) = \sqrt{1 - \cos^2\theta(T)} = \frac{\sqrt{4n - t^2}}{2\sqrt{n}}.$$

Also notice that the elliptic integral is invariant under $\theta \mapsto \pi - \theta$;

$$\int_{-\infty}^{\infty} \frac{e^{-2(\pi-\theta)r}}{1+e^{-2\pi r}} h(r)\, dr = \int_{-\infty}^{\infty} \frac{e^{2\theta r}}{1+e^{2\pi r}} h(r)\, dr$$
$$\left\{ r \leftarrow -r \right\} = \int_{-\infty}^{\infty} \frac{e^{-2\theta r}}{1+e^{-2\pi r}} h(r)\, dr.$$

Using all this, we see that (3.14) equals

(3.17)
$$\sum_{\substack{t\in\mathbb{Z} \\ t^2-4n>0 \\ \sqrt{t^2-4n}\notin\mathbb{Q}}} \sum_{T\in H_{n,t}//\Gamma_R} \frac{\ln \mathcal{N}(T_0)}{2\sqrt{t^2-4n}} \cdot g\left( \ln \frac{(|t| + \sqrt{t^2-4n})^2}{4n} \right)$$
$$+ \sum_{\substack{t\in\mathbb{Z} \\ t^2-4n<0}} \sum_{T\in H_{n,t}//\Gamma_R} \frac{1}{2M_T\sqrt{4n-t^2}} \cdot \int_{-\infty}^{\infty} \frac{e^{-2r\cdot \arccos(t/2\sqrt{n})}}{1+e^{-2\pi r}} \cdot h(r)\, dr.$$

Notice that we could restrict the sum over $t$ further, using Lemma 3.4 or Lemma 3.5, depending on which case we are in.

Now fix some $t \in \mathbb{Z}$ such that $\sqrt{t^2-4n} \notin \mathbb{Q}$, and such that $H_{n,t}$ is nonempty (cf. Lemma 3.4, Lemma 3.5). We will show how to enumerate the $\Gamma_R$-conjugacy classes in $H_{n,t}$, and for each conjugacy class, how to get hold of the number $\mathcal{N}(T_0)$ or $M_T$ occuring in (3.17). We first *fix* some $T_t \in H_{n,t}$. We also take $P(X)$ as in (3.2) and $d, l$ as in (3.3).

Now

$$P(T_t) = T_t^2 - (T_t + \overline{T_t}) \cdot T_t + \overline{T_t} T_t = 0,$$

and since $\sqrt{t^2-4n} \notin \mathbb{Q}$, $P(X)$ is irreducible, and the subalgebra $\mathbb{Q}[T_t] \subset B$ generated by $T_t$ over $\mathbb{Q}$ is actually a quadratic field, isomorphic to $\mathbb{Q}(\sqrt{d})$. We know that the restriction of $N_B$ and $\mathrm{tr}_B$ to $\mathbb{Q}[T_t]$ is simply the norm and the trace in the field $\mathbb{Q}[T_t]$ over $\mathbb{Q}$ [M2, p. 201(top)].

Recall the definition of $\mathfrak{r}[f]$, (3.4). Using $\mathbb{Q}[T_t] \cong \mathbb{Q}(\sqrt{d})$, we will from now on write $\mathfrak{r}[f]$ also for the corresponding order *in* $\mathbb{Q}[T_t]$ (of course, this does not depend on which of the two possible isomorphisms $\mathbb{Q}[T_t] \to \mathbb{Q}(\sqrt{d})$ we choose).

By [M2, Lemma 5.2.2(2)], the set of all elements in $B$ that have norm $n$ and trace $t$ equals

$$C(T_t) := \{\delta T_t \delta^{-1} \mid \delta \in B^\times\}. \tag{3.18}$$

In particular, $H_{n,t} = R \cap C(T_t)$.

For any $\delta \in B^\times$, $(\mathbb{Q}[T_t] \cap \delta^{-1} R \delta)$ is an order in $\mathbb{Q}[T_t]$, by (1.1). For any order $\mathfrak{r}$ in $\mathbb{Q}[T_t]$ we define:

$$C(T_t, \mathfrak{r}) = \left\{ \delta T_t \delta^{-1} \mid \delta \in B^\times, \quad \mathbb{Q}[T_t] \cap \delta^{-1} R \delta = \mathfrak{r} \right\}. \tag{3.19}$$

Our definitions (3.18) and (3.19) are the same as in [M2, pp. 244-245].

**Lemma 3.6.** *We have a disjoint union*

$$C(T_t) = \bigsqcup_{f=1}^{\infty} C(T_t, \mathfrak{r}[f]). \tag{3.20}$$

*Also,*

$$H_{n,t} = \bigsqcup_{f|l} C(T_t, \mathfrak{r}[f]). \tag{3.21}$$

*Each $C(T_t, \mathfrak{r})$ is closed under $\Gamma_R$-conjugation.*

*Proof.* Clearly $C(T_t) = \cup_{f=1}^{\infty} C(T_t, \mathfrak{r}[f])$, since there are no other orders than $\mathfrak{r}[1]$, $\mathfrak{r}[2], \ldots$ in $\mathbb{Q}[T_t]$. To prove disjointness, we must show that if $\delta T_t \delta^{-1} = \delta' T_t \delta'^{-1}$ ($\delta, \delta' \in B^\times$), then $\delta^{-1} R \delta$ and $\delta'^{-1} R \delta'$ have the same intersection with $\mathbb{Q}[T_t]$. But from $\delta T_t \delta^{-1} = \delta' T_t \delta'^{-1}$ we get $\delta^{-1}\delta' \in \mathbb{Q}[T_t]$ by [M2, Lemma 5.2.2(3)], and thus

$$\begin{aligned}\left(\mathbb{Q}[T_t] \cap \delta'^{-1} R \delta'\right) &= (\delta^{-1}\delta') \left(\mathbb{Q}[T_t] \cap \delta'^{-1} R \delta'\right) (\delta^{-1}\delta')^{-1} \\ &= \left(\mathbb{Q}[T_t] \cap \delta^{-1} R \delta\right).\end{aligned}$$

This proves (3.20).

Next, notice that for any order $\mathfrak{r}$ in $\mathbb{Q}[T_t]$, and any $\delta \in B^\times$ such that $\delta T_t \delta^{-1} \in C(T_t, \mathfrak{r})$, we have the following chain of equivalent statements:

$$\begin{aligned}T_t \in \mathfrak{r} &\Longleftrightarrow T_t \in \left(\mathbb{Q}[T_t] \cap \delta^{-1} R \delta\right) \Longleftrightarrow T_t \in \delta^{-1} R \delta \\ &\Longleftrightarrow \delta T_t \delta^{-1} \in R \Longleftrightarrow \delta T_t \delta^{-1} \in H_{n,t},\end{aligned}$$

since $N_B$ and $\mathrm{tr}_B$ are preserved under conjugation.

In other words:

$$\begin{aligned}&\text{if } T_t \in \mathfrak{r}, \quad \text{then } C(T_t, \mathfrak{r}) \subset H_{n,t}; \\ &\text{if } T_t \notin \mathfrak{r}, \quad \text{then } C(T_t, \mathfrak{r}) \cap H_{n,t} = \emptyset.\end{aligned}$$

But by (3.5), (3.6), the orders of $\mathfrak{r}$ in $\mathbb{Q}[T_t]$ that contains $T_t$ are exactly those in $\{\mathfrak{r}[f] \mid \quad f \mid l\}$. Now (3.21) follows from (3.20) and $H_{n,t} \subset C(T_t)$.

Finally, using $\beta^{-1}R\beta = R$ for any $\beta \in R^\times$, we also see that each $C(T_t, \mathfrak{r})$ is closed under $R^\times$-conjugation, and in particular under $\Gamma_R$-conjugation. $\square$

**Lemma 3.7.** *In the situation above, take $f \mid l$ and $T \in C(T_t, \mathfrak{r}[f]) \subset H_{n,t}$. If $T$ is elliptic, i.e. if $d < 0$, then $M_T$ in (3.17) equals:*

$$M_T = \frac{1}{2}\left|\mathfrak{r}[f]^1\right| = \frac{\left|\mathfrak{r}[1]^1\right|}{2\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}$$

*(cf. (3.7)). If $T$ is hyperbolic, i.e. if $d > 0$, then $\mathcal{N}(T_0)$ in (3.17) equals:*

$$\mathcal{N}(T_0) = \epsilon_d^{2\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]},$$

*where $\epsilon_d$ is the proper fundamental unit in $\mathbb{Q}(\sqrt{d})$.*

*Proof.* Choose some $\delta \in B^\times$ such that $T = \delta T_t \delta^{-1}$ and $\mathbb{Q}[T_t] \cap \delta^{-1} R \delta = \mathfrak{r}[f]$. Then $Z_{\Gamma_R}(T)$, the centralizer of $T$ in $\Gamma_R$ is by [M2, Lemma 5.2.2(3)]:

$$Z_{\Gamma_R}(T) = \mathbb{Q}[T] \cap \Gamma_R = \delta\left(\mathbb{Q}[T_t] \cap \delta^{-1}\Gamma_R\delta\right)\delta^{-1} = \delta \cdot \mathfrak{r}[f]^1 \cdot \delta^{-1}. \tag{3.22}$$

Hence if $d < 0$, then $M_T = \frac{1}{2}\left|Z_{\Gamma_R}(T)\right| = \frac{1}{2}\left|\mathfrak{r}[f]^1\right|$.

If $d > 0$, then by (3.8) and (3.22) we can take for $T_0$ the image of $\epsilon_d^{\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}$ under the composite of the two isomorphisms $\mathbb{Q}(\sqrt{d}) \cong \mathbb{Q}[T_t]$ and $\mathbb{Q}[T_t] \ni \beta \mapsto \delta\beta\delta^{-1} \in \mathbb{Q}[T]$. Then $\mathrm{tr}_B(T_0) = \mathrm{tr}\left(\epsilon_d^{\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}\right)$, and of course, $N_B(T_0) = 1$. Hence by Lemma 3.1 and Remark 3.2, $\mathcal{N}(T_0) = \epsilon_d^{2\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}$. $\square$

Using this, we now see that (3.17) can be written as

$$\sum_{\substack{t \in \mathbb{Z} \\ t^2 - 4n > 0 \\ \sqrt{t^2-4n} \notin \mathbb{Q}}}^{(*)} \sum_{f \mid l} \left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right| \cdot \frac{\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right] \cdot \ln \epsilon_d}{\sqrt{t^2 - 4n}} \cdot g\left(\ln \frac{(|t| + \sqrt{t^2 - 4n})^2}{4n}\right)$$
$$+ \sum_{\substack{t \in \mathbb{Z} \\ t^2 - 4n < 0}}^{(*)} \sum_{f \mid l} \left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right| \cdot \frac{\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]}{\left|\mathfrak{r}[1]^1\right| \cdot \sqrt{4n - t^2}} \cdot \int_{-\infty}^{\infty} \frac{e^{-2r \cdot \arccos(t/2\sqrt{n})}}{1 + e^{-2\pi r}} \cdot h(r)\, dr. \tag{3.23}$$

Here $\sum^{(*)}$ means that the sum should be restricted to those $t$ for which there exists some $T_t \in H_{n,t}$. This set of $t$:s is described in Lemma 3.4 (for case [I]), Lemma 3.5 (for case [II]). Also, $d, l$ are "implicit variables" determined by $t$ as in (3.3), and $\mathfrak{r}[f]$, $\mathfrak{r}[1]$ refers to orders in $\mathbb{Q}[T_t] \cong \mathbb{Q}(\sqrt{d})$.

Now take some $t, d, l, f$ occurring in one of the sums in (3.23). We have by [M2, Lemma 6.5.2][14], as on [M2, p.248(first half)]:

$$\left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right| = \mathbf{h}(\mathfrak{r}[f]) \cdot \prod_{v \in \{\infty,2,3,5,\ldots\}} \left|C_v(T_t, \mathfrak{r}[f])//R_v^\times\right|. \tag{3.24}$$

Here $\mathbf{h}(\mathfrak{r}[f])$ is the class number of $\mathfrak{r}[f]$, as defined above. For $v = p$ a prime, $R_p^\times$ is the group of invertible elements in the order $R_p$, while for $v = \infty$ we write $R_\infty^\times = \mathrm{GL}^+(2, \mathbb{R})$. Also:

$$C_p(T_t, \mathfrak{r}[f]) = \{xT_tx^{-1} \mid x \in B_p^\times, \quad \mathbb{Q}_p[T_t] \cap x^{-1}R_px = (\mathfrak{r}[f])_p\}$$
$$\text{for } p \text{ a prime number, and:}$$
$$C_\infty(T_t, \mathfrak{r}[f]) = \{xT_tx^{-1} \mid x \in B_\infty^\times\}.$$

One easily checks that each $C_v(T_t, \mathfrak{r}[f])$ is closed under $R_v^\times$-conjugation, cf. the proof of Lemma 3.6.

The factor in (3.24) corresponding to $v = \infty$ is easily determined, as in [M2, (6.6.1)]:

$$\left|C_\infty(T_t, \mathfrak{r}[f])//R_\infty^\times\right| = \begin{cases} 2 & \text{if } T_t \text{ is elliptic, i.e. } d < 0 \\ 1 & \text{if } T_t \text{ is hyperbolic, i.e. } d > 0 \end{cases}.$$

Now assume $v = p$, a prime. In case [I] we have $R_p = M_2(\mathbb{Z}_p)$ as soon as $p \nmid N$. In case [II], if $p \nmid d_B$, we have $B_p \cong M_2(\mathbb{Q}_p)$, and we may identify $B_p$ and $M_2(\mathbb{Q}_p)$ in a way such that $R_p = M_2(\mathbb{Z}_p)$, by [M2, Lemma 5.2.3(2)]. For such primes $p$, we have by [M2, Thm. 6.6.7]:

$$\text{if } R_p = M_2(\mathbb{Z}_p): \qquad \left|C_p(T_t, \mathfrak{r}[f])//R_p^\times\right| = 1.$$

In case [II] we are now left with those primes $p$ that divide $d_B$. For such a $p$, $B_p$ is a division algebra, and $R_p$ is its unique maximal order [M2, Lemma 5.2.3(1)], and we have by [M2, Thm. 6.6.8], (3.10), (3.11):

$$\text{Case [II], if } p \mid d_B: \qquad \left|C_p(T_t, \mathfrak{r}[f])//R_p^\times\right| = \begin{cases} 0 & \text{if } p \mid f \\ 1 & \text{if } p \nmid f, (d/p) = 0 \\ 2 & \text{if } p \nmid f, (d/p) = -1 \end{cases}$$

(The case $(d/p) = 1$ is impossible, by Lemma 3.5).

[14] Notice that the proof of Lemma 6.5.2 in [M2] goes through, word by word, also in the case $\alpha \in T^{h'}$ (Miyake's notation), that is, when $\alpha$ is hyperbolic with fixpoints that are *not* cusps.

Finally, in case [I] we are left with those primes $p$ that divide $N$. For such a $p$, a complete set of representatives for $C_p(T_t, \mathfrak{r}[f])//R_p^\times$ is given in [M2, Thm. 6.6.6][15]. We should take $\nu = 1$ therein (see [M2, (6.6.3)]), since $N$ is square free, and $\rho = \mathrm{ord}_p(l/f)$ (use (3.6) and (3.11)). Then notice that $t^2 - 4n = l^2 d$ is divisible by $p^{2\rho}$. The theorem (when corrected) tells us that

$$\left|C_p(T_t, \mathfrak{r}[f])//R_p^\times\right| = \left|\Omega/p^{1+\rho}\right| + \left|\Omega'/p^{1+\rho}\right|, \tag{3.25}$$

where

$$\begin{aligned} &\rho = \mathrm{ord}_p(l/f) \\ &\Omega = \{\xi \in \mathbb{Z}_p \mid \xi^2 - t\xi + n \equiv 0 \pmod{p^{1+2\rho}}\}; \\ &\Omega' = \begin{cases} \{\xi \in \mathbb{Z}_p \mid \xi^2 - t\xi + n \equiv 0 \pmod{p^{2+2\rho}}\} & \text{if } p^{2\rho+1} \mid l^2 d \\ \emptyset & \text{if } p^{2\rho+1} \nmid l^2 d \end{cases}. \end{aligned}$$

It is easily shown that both $\Omega$ and $\Omega'$ are closed under addition with any element from $p^{1+\rho}\mathbb{Z}_p$. (This will also be seen in the computations below.) In (3.25), $\Omega/p^{1+\rho}$ denotes a complete set of representatives for $\Omega \pmod{p^{1+\rho}\mathbb{Z}_p}$, and $\Omega'/p^{1+\rho}$ is the analogous set of representatives for $\Omega'$.

We now evaluate (3.25). Assume $p \neq 2$. For $\xi \in \mathbb{Z}_p$ we find

$$\begin{aligned} &\xi^2 - t\xi + n \equiv 0 \pmod{p^{1+2\rho}} \iff (2\xi - t)^2 \equiv l^2 d \pmod{p^{1+2\rho}} \\ &\iff \xi \equiv \frac{t}{2} \pmod{p^\rho} \quad \wedge \quad \left(\frac{2\xi - t}{p^\rho}\right)^2 \equiv (lp^{-\rho})^2 d \pmod{p}, \end{aligned} \tag{3.26}$$

since $p^{2\rho} \mid l^2$. (Notice that $t/2 \in \mathbb{Z}_p$, since $p \neq 2$.) First assume that $\rho < \mathrm{ord}_p l$ or $p \mid d$. Then $(lp^{-\rho})^2 d \equiv 0 \pmod{p}$, and the above equation is equivalent to $\xi \equiv t/2 \pmod{p^{\rho+1}}$. We also have, in this case, $p^{2\rho+1} \mid l^2 d$, and by manipulations very similar to (3.26) we find that (since $p^2 \nmid d$, by (3.1)):

$$\Omega' = \begin{cases} t/2 + p^{\rho+1}\mathbb{Z}_p & \text{if } \rho < \mathrm{ord}_p l \\ \emptyset & \text{if } \rho = \mathrm{ord}_p l,\ p \mid d \end{cases}.$$

Now assume that $\rho = \mathrm{ord}_p l$ and $p \nmid d$. Then (3.26) is seen to have two distinct solutions $\xi \pmod{p^{\rho+1}}$ if $(d/p) = 1$, but no solutions if

[15] There is a misprint in the statement of [M2, Thm. 6.6.6]: "$\Omega'/p^{\nu+\rho+1}$" should be replaced by $\Omega'/p^{\nu+\rho}$, i.e. a complete set of representatives of $\Omega'(\alpha, \mathfrak{r}_p) \bmod p^{\nu+\rho}$.

$(d/p) = -1$. Also, $\Omega' = \emptyset$, since $p^{2\rho+1} \nmid l^2 d$. We have shown that, for $p \neq 2$ :

$$\left|\Omega/p^{1+\rho}\right| + \left|\Omega'/p^{1+\rho}\right| = \begin{cases} 2 & \text{if } \rho < \text{ord}_p l \\ 1 & \text{if } \rho = \text{ord}_p l,\ p \mid d \\ 2 & \text{if } \rho = \text{ord}_p l,\ (d/p) = 1 \\ 0 & \text{if } \rho = \text{ord}_p l,\ (d/p) = -1 \end{cases}. \tag{3.27}$$

A slight modification of the analysis above shows that (3.27) is also valid for $p = 2$. We can collect our result as follows, using $\rho = \text{ord}_p(l/f)$ and referring to (3.25):

$$\text{Case [I], if } p \mid N: \qquad \left|C_p(T_t, \mathfrak{r}[f])//R_p^\times\right| = \begin{cases} 2 & \text{if } p \mid f \\ 1 + \left(\dfrac{d}{p}\right) & \text{if } p \nmid f \end{cases}.$$

We have now evaluated each "local factor" in (3.24). We get:

$$\text{Case [I]:} \quad \left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right| = \mathbf{h}(\mathfrak{r}[f]) \cdot \begin{Bmatrix} 2 & \text{if } d < 0 \\ 1 & \text{if } d > 0 \end{Bmatrix} \cdot \prod_{p \mid N} \begin{Bmatrix} 2 & \text{if } p \mid f \\ 1 + \left(\dfrac{d}{p}\right) & \text{if } p \nmid f \end{Bmatrix};$$

$$\text{Case [II]:} \quad \left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right| = \mathbf{h}(\mathfrak{r}[f]) \cdot \begin{Bmatrix} 2 & \text{if } d < 0 \\ 1 & \text{if } d > 0 \end{Bmatrix} \cdot \prod_{p \mid d_B} \begin{Bmatrix} 0 & \text{if } p \mid f \\ 1 - \left(\dfrac{d}{p}\right) & \text{if } p \nmid f \end{Bmatrix}.$$

Now recall Lemmas 3.4, 3.5. For any $t \in \mathbb{Z}$ with $\sqrt{t^2 - 4n} \notin \mathbb{Q}$ and $H_{n,t} = \emptyset$, we see that the above product, whichever applies, *vanishes* for all $f \mid l$. Hence when inserting our explicit expression for $\left|C(T_t, \mathfrak{r}[f])//\Gamma_R\right|$ in (3.23), we can remove the $(*)$-restriction on $t$.

We collect our result in two theorems.

**Definition 3.8.** *For $t \in \mathbb{Z}$, $n \in \mathbb{Z}^+$ such that $\sqrt{t^2 - 4n} \notin \mathbb{Q}$, we define*

$$A(t,n) := \begin{cases} \dfrac{\ln \epsilon_d}{\sqrt{t^2 - 4n}} \cdot g\left(\ln \dfrac{(|t| + \sqrt{t^2 - 4n})^2}{4n}\right) & \text{if } t^2 - 4n > 0 \\ \dfrac{2}{\left|\mathfrak{r}[1]^1\right| \cdot \sqrt{4n - t^2}} \cdot \displaystyle\int_{-\infty}^{\infty} \frac{e^{-2r \cdot arccos(t/2\sqrt{n})}}{1 + e^{-2\pi r}} \cdot h(r)\, dr & \text{if } t^2 - 4n < 0 \end{cases}.$$

*Here $d$ is as in* (3.3). *For $\left|\mathfrak{r}[1]^1\right|$ and $\epsilon_d$, cf. p. 21.*

**Theorem 3.9.** *Take $N, n \in \mathbb{Z}^+$, $N$ square free, $(N,n) = 1$. Then in the trace formula for the Hecke operator $T_n$ on $\Gamma_0(N)$ on p. 19, line* *three and four can be expressed as*

$$\sum_{\substack{t\in\mathbb{Z}\\ \sqrt{t^2-4n}\notin\mathbb{Q}}} \left( \sum_{f|l} \mathbf{h}(\mathfrak{r}[f]) \cdot \left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right] \cdot \prod_{p|N} \left\{ \begin{matrix} 2 & \text{if } p \mid f \\ 1+\left(\dfrac{d}{p}\right) & \text{if } p\nmid f \end{matrix} \right\} \right) \cdot A(t,n).$$

*Here* $d, l$ *are "implicit variables" depending on* $t$ *as in* (3.3)*, i.e.* $t^2-4n= l^2 d$, $l\in\mathbb{Z}^+$*, and* $d$ *is a fundamental discriminant* (3.1)*. Also,* $\mathfrak{r}[f]\subset \mathbb{Q}(\sqrt{d})$ *is as in* (3.4)*. For* $\mathbf{h}(\mathfrak{r}[f])$*, cf.* (3.12)*,* (3.13)*.*

*Remark* 3.10. The *hyperbolic part* of the sum in Theorem 3.9 can be shown to agree with [CL, Thm. 4.6] (when $g(u)$ is chosen as in [CL, (2.1)]). Also, for $N=n=1$, the enumeration of the hyperbolic conjugacy classes in $\Gamma_0(1)$ that is implicit in Theorem 3.9 agrees with [S1], [H4, pp. 514–518]. In fact, one can check that for $N=n=1$ the hyperbolic part of the sum in Theorem 3.9 equals

$$2\cdot \sum_{\substack{d\in\mathbb{Z}^+\\ d\equiv 0\vee d\equiv 1 \bmod 4\\ \sqrt{d}\notin\mathbb{Z}}} h(d)\cdot \sum_{k=1}^{\infty} \frac{\ln\epsilon_d}{\epsilon_d^k-\epsilon_d^{-k}}\cdot g(2k\cdot \ln\epsilon_d). \tag{3.28}$$

Here $h(d)$ is the class number for primitive quadratic forms as in [L1, Vol I, pp. 127–], and $\epsilon_d$ is as in [L1, Vol I, Satz 111]. For a simple comparison to [H4], take $g(u)=\exp(-(s-\frac{1}{2})|u|)$, and compare with [H4, p. 518, Prop. 2.8].

**Theorem 3.11.** *Let* $B$ *be an indefinite division quaternion algebra over* $\mathbb{Q}$*, and let* $R$ *be a maximal order in* $B$*. Let* $d_B>1$ *be the reduced discriminant of* $B$*. Take* $n\in\mathbb{Z}^+$*. Then the trace formula for the Hecke operator* $\widetilde{T}_n$ *(see* (1.11)*) on the cocompact group* $\Gamma_R$ *is:*

$$\sum_{\lambda\geq 0} h(r_\lambda)\cdot Tr\,(\widetilde{T}_n)_{|A(\lambda)}$$

$$= \left[\text{if } \sqrt{n}\in\mathbb{Z}: \qquad \frac{\prod_{p|d_B}(p-1)}{12\sqrt{n}}\cdot\int_{-\infty}^{\infty} r h(r)\tanh(\pi r)\,dr\right]$$

$$+\sum_{\substack{t\in\mathbb{Z}\\ \sqrt{t^2-4n}\notin\mathbb{Q}}} \left[\sum_{\substack{f|l\\ (f,d_B)=1}} \mathbf{h}(\mathfrak{r}[f])\cdot\left[\mathfrak{r}[1]^1:\mathfrak{r}[f]^1\right]\cdot\prod_{p|d_B}\left(1-\left(\frac{d}{p}\right)\right)\right]\cdot A(t,n).$$

*Here* $d, l$ *are "implicit variables" depending on* $t$ *as in* (3.3)*, i.e.* $t^2-4n= l^2 d$, $l\in\mathbb{Z}^+$*, and* $d$ *is a fundamental discriminant* (3.1)*. Also,* $\mathfrak{r}[f]\subset \mathbb{Q}(\sqrt{d})$ *is as in* (3.4)*. For* $\mathbf{h}(\mathfrak{r}[f])$*, cf.* (3.12)*,* (3.13)*.*

*Proof.* We only have to evaluate the identity term from (2.2). From (1.10) it is clear that

$$\left|\left\{\delta > 0 \mid \begin{pmatrix}\delta & 0\\ 0 & \delta\end{pmatrix} \in H_n\right\}\right| = \begin{cases}1 & \text{if } \sqrt{n} \in \mathbb{Z}\\ 0 & \text{otherwise}\end{cases}.$$

We also have

$$\mu(\Gamma_R \setminus \mathcal{H}) = \frac{\pi}{3} \cdot \prod_{p \mid d_B} (p-1), \tag{3.29}$$

cf. [E, pp. 49($q_1, q_2$),81]. This gives the stated result. □

## 4. Newforms

In this section we recall the formalism of oldspaces and newspaces (cf. Definitions 1.4, 1.5) in relation to the Hecke operators. The original reference for this is [AL]. Here we will follow the presentation in [M2, §4.5, §4.6] fairly closely. Both [AL] and [M2] deal with the holomorphic case of modular forms, but, as is well-known, the non-holomorphic case is very similar. (In principle, both the holomorphic and the non-holomorphic case should be possible to extract from the representation-theoretical results in [M1], [C1].)

The main results are collected in Theorem 4.6, see also Definition 4.5.

Recall the definition of $A(N, \lambda)$, (1.16). We will keep $\lambda$ (the eigenvalue) fixed and *positive*. We remind the reader that $A(N, \lambda)$ then only contains *cusp forms*. Also, $A(M, \lambda) \subset A(N, \lambda)$ when $M \mid N$. In this section, the level $N$ of the congruence group $\Gamma_0(N)$ will vary, and to obtain clear notation, we have to write ${}^N T_n$ instead of $T_n$ for the Hecke operator (2.9). Thus by (2.8), for $n \in \mathbb{Z}^+$, $f \in L_2(\Gamma_0(N) \setminus \mathcal{H})$ :

$$[{}^N T_n f](z) = \frac{1}{\sqrt{n}} \sum_{\substack{ad=n\\ (a,N)=1\\ a,d \in \mathbb{Z}^+}} \sum_{b \bmod d} f\left(\frac{az+b}{d}\right). \tag{4.1}$$

We know that ${}^N T_n$ maps $A(N, \lambda)$ into itself. When $(n, N) = 1$ we will sometimes continue to simply write $T_n$, for in this case we can remove the condition $(a, N) = 1$ in the right hand side of (4.1), which yields an expression that is independent of $N$.

Hecke introduced his operators to explain multiplicativity properties of the Fourier coefficients of certain modular forms [H1]. To see the nonholomorphic analog of this we need three lemmas. We also introduce some notation that will be useful throughout this section.

**Lemma 4.1.** *Take* $n, N \in \mathbb{Z}^+$, $\lambda = \frac{1}{4} + r^2 > 0$ *and* $f \in A(N, \lambda)$. *If the Fourier expansion of* $f$ *at* $i\infty$ *is*

$$f(z) = \sum_{m \neq 0} \alpha(m) \sqrt{y} K_{ir}(2\pi|m|y) e^{2\pi i m x} \qquad (z = x + iy), \tag{4.2}$$

*then*

$$[{}^N T_n f](z) = \sum_{m \neq 0} \left[ \sum_{\substack{a|(n,m),\\ (a,N)=1\\ a \in \mathbb{Z}^+}} \alpha(\frac{nm}{a^2}) \right] \sqrt{y} K_{ir}(2\pi|m|y) e^{2\pi i m x}.$$

*Proof.* This follows from (4.1) by a direct computation. Cf. [M2, Lemma 4.5.14]. □

In particular we have, for any $f \in A(N, \lambda)$, $n \in \mathbb{Z}^+$, $\tau \in \mathbb{C}$ :

$$ {}^N T_n f = \tau \cdot f \Longrightarrow \alpha(n) = \tau \cdot \alpha(1). \tag{4.3}$$

(By looking at the coefficient of index 1.)

**Lemma 4.2.** *Take* $N, n, m \in \mathbb{Z}^+$, *and let* $p$ *be a prime number.*

$$ {}^N T_n \circ {}^N T_m = {}^N T_m \circ {}^N T_n. \tag{4.4}$$

$$ {}^N T_n \circ {}^N T_m = {}^N T_{nm} \qquad \textit{if } (n,m) = 1. \tag{4.5}$$

$$ {}^N T_p \circ {}^N T_{p^e} = \begin{cases} {}^N T_{p^{e+1}} & \text{if } p \mid N \\ {}^N T_{p^{e+1}} + {}^N T_{p^{e-1}} & \text{if } p \nmid N \end{cases} \qquad \textit{for } e \geq 1. \tag{4.6}$$

*Proof.* Recall the proof of Lemma 1.7(b,c,d). This time, the element "$T(n)$" in the Hecke algebra "$\mathcal{R}(N)$" in [M2, (4.5.14)] is the same as our $H_n$ (2.7). Now our assertions follow from (2.9) and [M2, Thm. 4.5.3(1), Lemma 4.5.8(2), Lemma 4.5.7(2), (4.5.14):$T(p,p)$]. □

*Remark* 4.3. From (4.5) and (4.6) one easily deduces the general formula (cf. [M2, Lemma 4.5.12]), for any $n_1, n_2 \in \mathbb{Z}^+$ :

$$ {}^N T_{n_1} \circ {}^N T_{n_2} = \sum_{\substack{d|(n_1,n_2)\\ (d,N)=1}} {}^N T_{n_1 n_2 / d^2}. \tag{4.7}$$

For $f \in L_2(\Gamma_0(N) \setminus \mathcal{H})$ and $l \in \mathbb{Z}^+$ we define:

$$[\delta_l f](z) = \frac{1}{\sqrt{l}} f(lz), \qquad \text{as in (1.17).}$$

$$[Vf](z) = f(-\overline{z}). \tag{4.8}$$

$$[\omega_N f](z) = f\left(\frac{-1}{Nz}\right). \tag{4.9}$$

We say that $f$ is *even* if $Vf = f$ and *odd* if $Vf = -f$.

**Lemma 4.4.** *(a) $V$ is a bounded, self adjoint operator on $L_2(\Gamma_0(N) \setminus \mathcal{H})$, mapping $A(N, \lambda)$ into itself.*

*(b) $V \circ V = Id$.*

*(c) $V \circ {}^N T_n = {}^N T_n \circ V$ for all $n \in \mathbb{Z}^+$.*

*(d) If $f \in A(N, \lambda)$ has Fourier expansion ($z = x + iy$)*

$$f(z) = \sum_{m \neq 0} \alpha(m) \sqrt{y} K_{ir}(2\pi |m| y) e^{2\pi i m x},$$

*then*

$$[Vf](z) = \sum_{m \neq 0} \alpha(-m) \sqrt{y} K_{ir}(2\pi |m| y) e^{2\pi i m x};$$

$$[\delta_l f](z) = \sum_{m \neq 0} \alpha(m) \sqrt{y} K_{ir}(2\pi |lm| y) e^{2\pi i l m x}$$

*(e) $\delta_l$ is an injective linear map from $A(N, \lambda)$ into $A(lN, \lambda)$.*

*(f) $V \circ \delta_l = \delta_l \circ V : A(N, \lambda) \to A(lN, \lambda)$.*

*(g) $T_n \circ \delta_l = \delta_l \circ T_n : A(N, \lambda) \to A(lN, \lambda)$, when $(n, lN) = 1$.*

*(h) $\omega_N$ is the modular correspondence coming from the double coset $\Gamma_0(N) \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) \Gamma_0(N)$. In particular, $\omega_N$ maps $A(N, \lambda)$ into itself.*

*(i) $\omega_N$ is self adjoint on $L_2(\Gamma_0(N) \setminus \mathcal{H})$.*

*(j) $\omega_N \circ \omega_N = Id$.*

*(k) $({}^N T_n)^* = \omega_N \circ {}^N T_n \circ \omega_N$ on $L_2(\Gamma_0(N) \setminus \mathcal{H})$, for each $n \in \mathbb{Z}^+$. If $(n, N) = 1$, then $({}^N T_n)^* = {}^N T_n$ (and hence $\omega_N \circ {}^N T_n = {}^N T_n \circ \omega_N$).*

*(l) $\omega_N \circ \delta_l = \sqrt{\frac{N}{M}} \cdot \frac{1}{l} \cdot \delta_{N/(Ml)} \circ \omega_M : A(M, \lambda) \to A(N, \lambda)$, if $lM \mid N$.*

*(m) $V \circ \omega_N = \omega_N \circ V$.*

*Proof.* (a),(b),(c),(d) are proved by direct computations from the definitions. (For (a), one uses that if $\mathcal{F}$ is a fundamental domain for $\Gamma_0(N)$, then so is $-\overline{\mathcal{F}}$, and that $V$ commutes with the Laplace-Beltrami operator $D$ on $C^\infty(\Gamma_0(N) \setminus \mathcal{H})$. For (c), one uses (4.1).)

For (e), notice that if $f \in A(N,\lambda)$, then $\delta_l f \in A(\Gamma',\lambda)$, where $\Gamma' = \left(\begin{smallmatrix} l^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right) \Gamma_0(N) \left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right)$. But $\Gamma'$ contains $\Gamma_0(lN)$ as a subgroup of finite index, hence $\delta_l f \in A(lN,\lambda)$. Injectivity is obvious.

Now (f) is trivial, and (g) follows from (4.1) by a quick computation.

For (h), we need only verify that $\Gamma_0(N) \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) \Gamma_0(N) = \Gamma_0(N) \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. This follows from

$$\left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) \Gamma_0(N) \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)^{-1} = \Gamma_0(N), \tag{4.10}$$

which can be verified directly. Now (i) follows from $\left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)^{-1} = -\frac{1}{N}\left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$ and [S5, Lemma 2.2]. (j),(l),(m) are proved by quick computations.

For (k), notice that by (2.9) and [S5, Lemma 2.2],

$$({}^N T_n)^* = \frac{1}{\sqrt{n}} \left[\text{The modular correspondence coming from } (H_n)^{-1}\right]. \tag{4.11}$$

By a quick computation, using (2.7), one verifies that $(H_n)^{-1} = (Nn)^{-1} \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) H_n \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. Hence if $H_n = \bigsqcup_{i=1}^d \Gamma_0(N)\alpha_i$, we have $(H_n)^{-1} = \bigsqcup_{i=1}^d (Nn)^{-1} \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) \Gamma_0(N)\alpha_i \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. Using (4.10), we get $(H_n)^{-1} = \bigsqcup_{i=1}^d (Nn)^{-1}\Gamma_0(N) \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right) \alpha_i \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. Together with (4.11), this gives $({}^N T_n)^* = \omega_N \circ {}^N T_n \circ \omega_N$. Finally, if $(n,N) = 1$ we have $(H_n)^{-1} = n^{-1} H_n$ (as on p. 24), and thus $({}^N T_n)^* = {}^N T_n$. $\square$

Keep $N \in \mathbb{Z}^+$ and $\lambda > 0$ fixed. We define

$$\mathsf{S} := \{n \in \mathbb{Z}^+ \mid (n,N) = 1\}.$$

Now ${}^N T_n$ for all $n \in \mathsf{S}$, and $V$, form a family of commuting self adjoint operators on the finite dimensional space $A(N,\lambda)$, by (4.4) and Lemma 4.4(a,c,k). Hence there is a simultaneous diagonalization. For any $\epsilon \in \mathbb{C}$ and any sequence $\{\tau(n)\}_{n\in\mathsf{S}} \subset \mathbb{C}$, we define:

$$U\langle \epsilon, \{\tau\}_{\mathsf{S}}\rangle = \Big\{ f \in A(N,\lambda) \mid Vf = \epsilon f,\ {}^N T_n f = \tau(n) f,\ \forall n \in \mathsf{S} \Big\}. \tag{4.12}$$

The result is that $A(N,\lambda)$ is a direct orthogonal sum of certain "U-spaces" $U\langle \epsilon, \{\tau\}_{\mathsf{S}}\rangle$. Of course, any two different choices of $\langle \epsilon, \{\tau\}_{\mathsf{S}}\rangle$ yield orthogonal $U$-spaces, since all operators in (4.12) are self-adjoint. Also, if $U\langle \epsilon, \{\tau\}_{\mathsf{S}}\rangle \neq \{0\}$, then $\epsilon \in \{-1, 1\}$ and $\{\tau(n)\}_{n\in\mathsf{S}} \subset \mathbb{R}$ (by self-adjointness and Lemma 4.4(b)).

Take $f \in U\langle \epsilon, \{\tau\}_{\mathsf{S}}\rangle$, $f \neq 0$, and let its Fourier expansion be as in (4.2). Now *if* $\alpha(1) \neq 0$ we may renormalize $f$ so that $\alpha(1) = 1$. Then from (4.3) and Lemma 4.4(d):

$$\alpha(n) = \tau(n), \qquad \alpha(-n) = \epsilon \cdot \tau(n) \quad \text{for all } n \in \mathsf{S},$$

and from (4.7) we then obtain

$$\alpha(n_1)\alpha(n_2) = \sum_{d|(n_1,n_2)} \alpha(\frac{n_1 n_2}{d^2}) \qquad \text{for all } n_1, n_2 \in \mathsf{S}.$$

This is the Hecke type multiplicativity.

Let us note that if $f \in U\langle \epsilon, \{\tau\}_{\mathsf{S}} \rangle$ with $\alpha(1) = 1$ happen to be an eigenfunction to *all* Hecke operators ${}^N T_n$ ($n \in \mathbb{Z}^+$), then by (4.5), (4.6), the Fourier coefficients will satisfy (for $n, m, e \in \mathbb{Z}^+$, $p$ a prime):

$$\begin{cases} \alpha(-n) = \epsilon\alpha(n) & \\ \alpha(nm) = \alpha(n)\alpha(m) & \text{if } (n,m) = 1 \\ \alpha(p^{e+1}) = \alpha(p)\alpha(p^e) & \text{if } p \mid N. \\ \alpha(p^{e+1}) = \alpha(p)\alpha(p^e) - \alpha(p^{e-1}) & \text{if } p \nmid N. \end{cases} \tag{4.13}$$

*If we restrict ourselves to the newspace* $A_{\mathrm{new}}(N, \lambda) \subset A(N, \lambda)$ (cf. Definition 1.5), then *there is* a basis of functions which have these nice properties. This fact follows from parts of Theorem 4.6 below, which collects all the main results to be proved in this section. Before we state this theorem, recall Definitions 1.4 and 1.5. We also give one more definition:

**Definition 4.5.** *$f \in A_{new}(N, \lambda)$ is called a* newform *if*

*(i) $f$ is an eigenfunction of $V$ and all ${}^N T_n$ with $(n, N) = 1$.*

*(ii) in the Fourier expansion (4.2), $\alpha(1) = 1$.*

Hence if $f \in A_{\mathrm{new}}(N, \lambda)$ is a newform with Fourier expansion (4.2), then $f \in U\langle \alpha(-1), \{\alpha\}_{\mathsf{S}} \rangle$, by (4.3) and Lemma 4.4(d).

**Theorem 4.6.** *Fix $\lambda > 0$ and keep $N, M \in \mathbb{Z}^+$.*

*(a) There are only finitely many newforms in $A_{new}(N, \lambda)$. Enlist them as $f_1^{(N)}, \ldots, f_{m_N}^{(N)}$. They form an orthogonal basis in $A_{new}(N, \lambda)$.*

*(b) Each $f_i^{(N)}$ is an eigenfunction with real eigenvalues to ${}^N T_n$ and $({}^N T_n)^*$ for* all $n \geq 1$, *and to $\omega_N$.*

*(c) A basis in $A(N, \lambda)$ is given by the elements in*

$$\{\delta_l f_i^{(M)} \mid \quad l, M \in \mathbb{Z}^+,\ lM \mid N,\ 1 \leq i \leq m_M\}.$$

*(d) Take $\lambda_1, \lambda_2 > 0$. If $f_1 \in A_{new}(N, \lambda_1)$ and $f_2 \in A_{new}(M, \lambda_2)$ are newforms, then either $f_1 \equiv f_2$, $\lambda_1 = \lambda_2$ and $N = M$, or else there are infinitely many primes $p$ ($p \nmid NM$) such that $f_1, f_2$ have different eigenvalues with respect to $T_p$. (Concerning notation, see the remark following* (4.1)*.)*

Cf. [AL, Thm. 5] and [M2, Thm. 4.6.13, Thm. 4.6.19]. Our case differs from the holomorphic case mainly on three points:

I)We have to keep track of even, odd eigenfunctions with the operator $V$.

II)The functional equation for the $L-$functions is different.

III)The (analog of the) Ramanujan-Petersson conjecture is not proved in our case, but we still have access to estimates of the Fourier coefficients that are good enough.

The remaining part of this section is devoted to the proof of Theorem 4.6. We keep the eigenvalue $\lambda > 0$ fixed.

**Lemma 4.7.** *Take $L, N \in \mathbb{Z}^+$ and $f \in A(N, \lambda)$ with Fourier expansion as in* (4.2)*. Assume that $\alpha(n) = 0$ for each $n \in \mathbb{Z}$ relatively prime to $L$. Then there are functions $f_p \in A(N/p, \lambda)$ (for primes $p$ dividing $(L, N)$) such that*

$$f \equiv \sum_{p|(L,N)} \delta_p f_p.$$

*In particular, if $(L, N) = 1$, then $f \equiv 0$.*

*Proof.* This is the exact analog of [M2, Thm. 4.6.8, $(\chi \equiv 1)$], and one can check that the proof (which includes the lemmas on p.155-159) carries over, with only simple and obvious changes. □

**Corollary 4.8.** *Take $L, N \in \mathbb{Z}^+$ and $f \in A_{new}(N, \lambda)$ with Fourier expansion as in* (4.2)*. If $f$ is an eigenfunction ($\not\equiv 0$) of $V$ and ${}^N T_n$ for all $n$ relatively prime to $L$, then $\alpha(1) \neq 0$.*

*Proof.* If $\alpha(1) = 0$, then by (4.3) and Lemma 4.4(d), $\alpha(n) = 0$ for all $n \in \mathbb{Z}$ relatively prime to $L$, and hence by Lemma 4.7, $f \in A_{\text{old}}(N, \lambda)$, a contradiction. □

Now certain parts of Theorem 4.6 are easily proved. Notice that if $n \in \mathsf{S}$, then $A_{\text{old}}(N, \lambda)$ is stable under the operator ${}^N T_n$, for if $lM \mid N$, $M \neq N$ and $f \in A(M, \lambda)$, then

$${}^N T_n[\delta_l f] = {}^{lM} T_n[\delta_l f] = \delta_l[{}^M T_n f] \in \delta_l A(M, \lambda) \subset A_{\text{old}}(N, \lambda),$$

by the remark following (4.1) and Lemma 4.4(g). $A_{\text{old}}(N, \lambda)$ is also stable under $V$, by Lemma 4.4(f). Since these operators are self adjoint on $A(N, \lambda)$, we get from Definition 1.5:

(4.14) $\quad A_{\text{new}}(N, \lambda)$ is stable under $V$ and ${}^N T_n$ for all $n \in \mathsf{S}$.

Hence we can diagonalize $A_{\text{new}}(N, \lambda)$ with respect to $V$ and all ${}^N T_n$ ($n \in \mathsf{S}$), and after renormalization, using Cor. 4.8 (with $L = N$), we obtain a basis in $A_{\text{new}}(N,\lambda)$ consisting of newforms. Recall our remark after Definition 4.5. Notice, that if $f$ and $g$ are *any* two newforms in $A_{\text{new}}(N,\lambda)$ belonging to the same $U$-space, then Cor. 4.8 applied to $f-g$ yields $f\equiv g$. Hence any two *different* newforms in $A_{\text{new}}(N,\lambda)$ belong to different $U$-spaces, and are thus orthogonal. This proves Theorem 4.6(a).

□

We now also obtain part of Theorem 4.6(c):

$$\{\delta_l f_i^{(M)} \mid \quad lM \mid N,\ 1\le i\le m_M\} \quad \text{spans } A(N,\lambda). \tag{4.15}$$

This follows easily by complete induction over $N$, using Definitions 1.4 and 1.5 and Theorem 4.6(a).

The next lemma is the analog of [M2, Thm. 4.6.12].

**Lemma 4.9.** *Take $L,N\in\mathbb{Z}^+$, $f\in A_{new}(N,\lambda)$ and $g\in A(N,\lambda)$. If $f$ and $g$ are (nonzero) eigenfunctions with the same eigenvalue to $V$ and ${}^NT_n$ for all $n\in\mathbb{Z}^+$ relatively prime to $L$, then $g$ is a constant multiple of $f$ (which in turn is a constant multiple of a newform in $A_{new}(N,\lambda)$).*

*Proof.* We may replace $L$ by some multiple of it, and we may thus assume $N\mid L$. Let $\mathsf{S}'=\{n\in\mathbb{Z}^+\mid (n,L)=1\}$, and define $U\langle\epsilon,\{\tau\}_{\mathsf{S}'}\rangle$ in analogy with (4.12). Since $N\mid L$, $({}^NT_n)^*={}^NT_n$ for all $n\in\mathsf{S}'$, hence distinct choices of $\langle\epsilon,\{\tau\}_{\mathsf{S}'}\rangle$ yield orthogonal $U$-spaces.

Let $f$ have Fourier expansion as in (4.2). By Cor. 4.8, $\alpha(1)\neq 0$, and we may assume $\alpha(1)=1$. Then by (4.3) and Lemma 4.4(d), $f,g\in U\langle\alpha(-1),\{\alpha\}_{\mathsf{S}'}\rangle$. We will prove the lemma by showing that $U\langle\alpha(-1),\{\alpha\}_{\mathsf{S}'}\rangle$ is a one-dimensional space spanned by one newform in $A_{\text{new}}(N,\lambda)$.

Take $f_i^{(M)}$ as in (4.15). $f_i^{(M)}$ is a newform on $\Gamma_0(M)$ ($M\mid N$), so by the remark following (4.1), $f_i^{(M)}$ belongs to some space $U\langle\epsilon,\{\tau\}_{\mathsf{S}'}\rangle$, and by Lemma 4.4(f,g), $\delta_l f_i^{(M)}$ belongs to the very same space for each $l\mid\dfrac{N}{M}$. Since distinct $\langle\epsilon,\{\tau\}_{\mathsf{S}'}\rangle$'s yield orthogonal $U$-spaces, it now suffices to prove that $f_i^{(M)}\notin U\langle\alpha(-1),\{\alpha\}_{\mathsf{S}'}\rangle$ when $M<N$, and that $f_i^{(N)}\in U\langle\alpha(-1),\{\alpha\}_{\mathsf{S}'}\rangle$ for *exactly one* $i\in\{1,...,m_N\}$.

But *if* $f_i^{(M)}\in U\langle\alpha(-1),\{\alpha\}_{\mathsf{S}'}\rangle$, then as in the proof of Cor. 4.8 we get $f-f_i^{(M)}\in A_{\text{old}}(N,\lambda)$. In the case $M<N$ this yields the contradiction $f\in A_{\text{old}}(\lambda,N)$. In the case $M=N$, the only possibility is $f_i^{(N)}\equiv f$. Clearly there must exist *some* $i\in\{1,...,m_N\}$ for which

$f_i^{(N)} \in U\langle\alpha(-1), \{\alpha\}_{\mathsf{S}'}\rangle$, otherwise $f$ would be orthogonal to all of $A_{\mathrm{new}}(\lambda, N)$. $\square$

Now we can prove Theorem 4.6(b): Take $f = f_i^{(N)}$ for some $i \in \{1, ..., m_N\}$; then $f \in U\langle\epsilon, \{\tau\}_{\mathsf{S}}\rangle$ for some $\langle\epsilon, \{\tau\}_{\mathsf{S}}\rangle$. Take *an arbitrary* $m \in \mathbb{Z}^+$ and look at ${}^N T_m f$. Using (4.4) and Lemma 4.4(c), we get

$$V[{}^N T_m f] = \epsilon \cdot {}^N T_m f \quad \text{and}$$
$${}^N T_n[{}^N T_m f] = \tau(n) \cdot {}^N T_m f, \quad \forall n \in \mathsf{S}.$$

Hence by Lemma 4.9($L = N$), ${}^N T_m f$ is a constant multiple of $f$; ${}^N T_m f = \mu f$ for some $\mu \in \mathbb{C}$. The same argument applies to $\omega_N f$, by Lemma 4.4(k,m), hence $\omega_N f = \eta f$, where $\eta \in \{-1, 1\}$ by Lemma 4.4(j).

Now by Lemma 4.4(k), $({}^N T_m)^* f = \eta\mu\eta f = \mu f$, and thus $\mu\langle f, f\rangle = \langle {}^N T_m f, f\rangle = \langle f, ({}^N T_m)^* f\rangle = \overline{\mu}\langle f, f\rangle$, i.e. $\mu \in \mathbb{R}$.

Hence Theorem 4.6(b) is proved. In particular we know that for any newform in $A_{\mathrm{new}}(N, \lambda)$, the Fourier coefficients satisfy (4.13), and are all real.

$\square$

Concerning Theorem 4.6(c): Notice that, by (4.15), $A(N, \lambda)$ is the sum of the spaces

$$W_{M,i} := \mathrm{span}_{\mathbb{C}} \left\{ \delta_l f_i^{(M)} \mid \quad l \mid \frac{N}{M} \right\} \qquad (M \mid N, \quad 1 \le i \le m_M).$$

By looking at the Fourier expansions (using Lemma 4.4(d) and Definition 4.5(ii)) we see that $\{\delta_l f_i^{(M)} \mid \quad l \mid \dfrac{N}{M}\}$ is a basis in $W_{M,i}$, for any $M \mid N$, $1 \le i \le m_M$. As in the proof of Lemma 4.9 (taking $L = N$), each space $W_{M,i}$ is a subspace of some $U\langle\epsilon, \{\tau\}_{\mathsf{S}}\rangle$. *Assuming* part (d) of Theorem 4.6, the $W_{M,i}$'s will lie in pairwise different $U\langle\epsilon, \{\tau\}_{\mathsf{S}}\rangle$-spaces, and will thus be pairwise orthogonal. This proves Theorem 4.6(c).

$\square$

Hence it only remains to prove (d) in Theorem 4.6.

We remark that if $\lambda_1 = \lambda_2$, $M \mid N$ (or $N \mid M$), and $f_1$, $f_2$ have the same eigenvalue with respect to $V$, then the assertion of (d) follows from Lemma 4.9 and (4.5), (4.6). In the general case however, there seems to be no shorter way known than to use $L$-functions and nontrivial estimates of the Fourier coefficients. Cf. [AL, p. 151(lines 12-14)]. We will proceed along this path.

**Lemma 4.10.** *Take $f \in A(N,\lambda)$, with Fourier expansion as in* (4.2)*. Then $\alpha(m) = O(|m|^{2/5+\epsilon})$ as $|m| \to \infty$.*

*Proof.* We refer to [P1, Cor. 1], where the result is obtained for an arbitrary cofinite Fuchsian group[16]. In our application below, any exponent better that $1/2$ is sufficient. □

**Lemma 4.11.** *If $f$ is a newform in $A_{new}(N,\lambda)$, with Fourier expansion as in* (4.2)*, and $p$ is a prime divisor of $N$, then*

$$\alpha(p) = 0 \quad \text{if } p^2 \mid N$$
$$\alpha(p) = \pm\frac{1}{\sqrt{p}} \quad \text{if } p^2 \nmid N.$$

*Proof.* This is the analog of [M2, Thm. 4.6.17(2,3), $\chi \equiv 1$], and the proof can be collected by mimicing (parts of) [M2, pp. 166-172]. □

**Lemma 4.12.** *Take $f \in A(N,\lambda)$ such that $Vf = \epsilon f$, ($\epsilon = \pm 1$), with Fourier expansion* (4.2)*, and form*

$$L(s,f) = \sum_{m=1}^{\infty} \alpha(m) \cdot m^{-s}.$$

*This Dirichlet series is absolutely convergent for* $\operatorname{Re} s > 1$*, and has an analytic continuation to an entire function. It satisfies the functional equation*

$$G_{N,\epsilon,r}(s) \cdot L(s,f) = \epsilon \cdot G_{N,\epsilon,r}(1-s) \cdot L(1-s, \omega_N f), \qquad \textit{where}$$

$$G_{N,\epsilon,r}(s) = \begin{cases} \pi^{-s} \cdot N^{s/2} \cdot \Gamma(\frac{s+ir}{2})\Gamma(\frac{s-ir}{2}) & \text{if } \epsilon = 1 \\ \pi^{-s} \cdot N^{s/2} \cdot \Gamma(\frac{s+1+ir}{2})\Gamma(\frac{s+1-ir}{2}) & \text{if } \epsilon = -1 \end{cases},$$

*and where $\omega_N f$ is as in* (4.9)*, and $\lambda = \frac{1}{4} + r^2 > 0$ as usual.*

*Proof.* This is well-known. The absolute convergence is proved using a standard Rankin-Selberg estimate for $\sum_{n \le X} |\alpha(n)|^2$; see [I, Thm. 3.2]. To prove analytic continuation and the functional equation, mimic [B, pp. 107–108(top)][17], where this is done for $N = 1$. □

---

[16]This has recently been improved; see [P2] (exponent $3/8+\epsilon$) and [BR] (exponent $1/3+\epsilon$). In fact, for congruence subgroups such as $\Gamma_0(N)$, much better results are known, see eg. [BDHI] (exponent $5/28+\epsilon$).

[17]Notice the misprint on p.107, line 6: "-1" should be "1"

Notice that if $f$ is a newform in $A_{\text{new}}(N, \lambda)$, then by expressing the last two lines of (4.13) in terms of generating functions, we get:

$$\sum_{m=0}^{\infty} \alpha(p^m) \cdot z^m = \frac{1}{1 - \alpha(p)z} \quad \text{if } p \mid N; \tag{4.16}$$

$$\sum_{m=0}^{\infty} \alpha(p^m) \cdot z^m = \frac{1}{1 - \alpha(p)z + z^2} \quad \text{if } p \nmid N. \tag{4.17}$$

By Lemma 4.11, the radius of convergence (for $z$) in (4.16) is $\geq \sqrt{p}$ and in (4.17) it is $\geq p^{-2/5}$, by Lemma 4.10. In particular, in (4.17), the zeros $z_1, z_2$ of $1 - \alpha(p)z + z^2$ satisfy

$$p^{-2/5} \leq |z_1|, |z_2| \leq p^{2/5}. \tag{4.18}$$

(Since $z_1 z_2 = 1$.)

Substituting $z = p^{-s}$ in (4.16), (4.17), and taking the product over all primes $p$, we get by (4.13) (line two) and Lemma 4.12:

$$L(s, f) = \prod_{p \mid N} \frac{1}{1 - \alpha(p)p^{-s}} \cdot \prod_{p \nmid N} \frac{1}{1 - \alpha(p)p^{-s} + p^{-2s}}, \tag{4.19}$$

with absolute convergence for $\operatorname{Re} s > 1$.

We are now in a position to finally prove Theorem 4.6(d). We take $f_1$ and $f_2$ to have Fourier coefficients $\{\alpha(n)\}_{n \in \mathbb{Z}}$ and $\{\beta(n)\}_{n \in \mathbb{Z}}$, respectively. Cf. (4.2), (4.3). Then ${}^N T_n f_1 = \alpha(n) f_1$, ${}^N T_n f_2 = \beta(n) f_2$ for $n = 1, 2, \ldots$ We assume that there is some $L \in \mathbb{Z}^+$, divisible by $NM$, such that $\alpha(p) = \beta(p)$ for all primes $p$ not dividing $L$, and we will then prove that $N = M$ and that $f_1 \equiv f_2$.

By (4.19) we have

$$\frac{L(s, f_1)}{L(s, f_2)} = \prod_{p \mid L} \frac{1 - \beta(p)p^{-s} + \delta_{p \nmid M} \cdot p^{-2s}}{1 - \alpha(p)p^{-s} + \delta_{p \nmid N} \cdot p^{-2s}}, \qquad \text{where} \tag{4.20}$$

$$\delta_{p \nmid N} = \begin{cases} 0 & \text{if } p \mid N \\ 1 & \text{if } p \nmid N \end{cases}.$$

At first, (4.20) holds for $\operatorname{Re} s > 1$, but by analytic continuation, using Lemma 4.12, it extends to an equality between meromorphic functions in the whole of $\mathbb{C}$.

By Definition 4.5(i), and Theorem 4.6(b), $f_1$ and $f_2$ are eigenfunctions of $V$ and of $\omega_N$ and $\omega_M$, respectively. Hence by Lemma 4.4(b,j),

$$\omega_N f_1 = \eta_1 f_1, \quad \omega_M f_2 = \eta_2 f_2, \quad V f_j = \epsilon_j f_j \quad (j = 1, 2)$$

for some $\eta_1, \epsilon_1, \eta_2, \epsilon_2 \in \{-1, 1\}$.

We write $\lambda_j = \frac{1}{4} + r_j^2 > 0$, $j = 1, 2$ as usual. By (4.20) and Lemma 4.12 we have

$$\begin{aligned}
&\prod_{p|L} \frac{1-\beta(p)p^{-s}+\delta_{p\nmid M}\cdot p^{-2s}}{1-\alpha(p)p^{-s}+\delta_{p\nmid N}\cdot p^{-2s}} \cdot \prod_{p|L} \frac{1-\alpha(p)p^{s-1}+\delta_{p\nmid N}\cdot p^{2s-2}}{1-\beta(p)p^{s-1}+\delta_{p\nmid M}\cdot p^{2s-2}} \\
&= \frac{L(s,f_1)}{L(s,f_2)} \cdot \frac{L(1-s,f_2)}{L(1-s,f_1)} \\
&= \eta_1\eta_2 \frac{L(s,f_1)}{L(s,f_2)} \cdot \frac{L(1-s,\omega_M f_2)}{L(1-s,\omega_N f_1)} \\
&= \eta_1\eta_2\epsilon_1\epsilon_2 \frac{G_{N,\epsilon_1,r_1}(1-s)}{G_{N,\epsilon_1,r_1}(s)} \cdot \frac{G_{M,\epsilon_2,r_2}(s)}{G_{M,\epsilon_2,r_2}(1-s)}. \qquad (4.21)
\end{aligned}$$

From the first expression, we see that this function is analytic outside some strip $|\operatorname{Re} s| < A$. On the other hand, in the last expression, none of the four "$G$-factors" has any zeros, and $G_{M,\epsilon_2,r_2}(1-s)$ has no poles for $\operatorname{Re} s < 1 - |\operatorname{Re}(ir_2)|$, while $G_{M,\epsilon_2,r_2}(s)$ has poles (double if $r_2 = 0$, otherwise simple) in each of the points

$$\begin{array}{ll}
\text{If } \epsilon_2 = 1: & \pm ir_2,\ -2 \pm ir_2,\ -4 \pm ir_2,\ -6 \pm ir_2, \ldots \\
\text{If } \epsilon_2 = -1: & -1 \pm ir_2,\ -3 \pm ir_2,\ -5 \pm ir_2,\ -7 \pm ir_2, \ldots
\end{array}$$

Hence infinitely many of these poles have to cancel against the poles of $G_{N,\epsilon_1,r_1}(s)$. But $G_{N,\epsilon_1,r_1}(s)$ has poles only in the points

$$\begin{array}{ll}
\text{If } \epsilon_1 = 1: & \pm ir_1,\ -2 \pm ir_1,\ -4 \pm ir_1,\ -6 \pm ir_1, \ldots \\
\text{If } \epsilon_1 = -1: & -1 \pm ir_1,\ -3 \pm ir_1,\ -5 \pm ir_1,\ -7 \pm ir_1, \ldots
\end{array}$$

Using $ir_1, ir_2 \in [0, \frac{1}{2}) \cup i[0, \infty)$, we see that the only possibility is $\epsilon_1 = \epsilon_2$ and $r_1 = r_2$. The equality (4.21) can now be written

$$\begin{aligned}
&\prod_{p|L} \frac{1-\beta(p)p^{-s}+\delta_{p\nmid M}\cdot p^{-2s}}{1-\alpha(p)p^{-s}+\delta_{p\nmid N}\cdot p^{-2s}} \cdot \prod_{p|L} \frac{1-\alpha(p)p^{s-1}+\delta_{p\nmid N}\cdot p^{2s-2}}{1-\beta(p)p^{s-1}+\delta_{p\nmid M}\cdot p^{2s-2}} \qquad (4.22) \\
&= \eta_1\eta_2 \left(\frac{N}{M}\right)^{1/2-s}.
\end{aligned}$$

Now mimic [M2, pp. 173(bottom)-174], using (4.18) and Lemmas 4.9, 4.11. We get $M = N$ and $f_1 \equiv f_2$, as asserted.

This completes the proof of Theorem 4.6. □

## 5. The trace formula for newforms

We now want to find the trace formula for Hecke operators on the *newspaces* of $\Gamma_0(N)$. From Theorem 4.6(a,c), it is evident that this can be expressed as a suitable linear combination – like a "sieve" – of the trace formulas for all congruence levels dividing $N$.

We start by introducing some arithmetic functions that enter into the computations. First recall that any function can be represented by a Dirichlet series in at most one way ([M2, Lemma 3.2.1] or [T1, 9.6]). The function $\beta(n)$ ($n \in \mathbb{Z}^+$) is now defined through the Dirichlet series

$$\zeta(s)^{-2} = \sum_{n=1}^{\infty} \frac{\beta(n)}{n^s}, \qquad \operatorname{Re} s > 1.$$

By squaring the relation $\zeta(s)^{-1} = \sum_{n=1}^{\infty} \mu(n) n^{-s} = \prod_p (1 - p^{-s})$, we see that

$$\beta(n) = \sum_{e|n} \mu(e) \mu\left(\frac{n}{e}\right) = \prod_{p|n} \left\{ \begin{matrix} -2 & \text{if } \operatorname{ord}_p(n) = 1 \\ 1 & \text{if } \operatorname{ord}_p(n) = 2 \\ 0 & \text{if } \operatorname{ord}_p(n) \geq 3 \end{matrix} \right\}.$$

As in Section 2 we let $d(n)$ denote the divisor function. Thus

$$d(n) := \sum_{e|n} 1 \qquad (n \in \mathbb{Z}^+); \qquad \zeta(s)^2 = \sum_{n=1}^{\infty} d(n) n^{-s}. \tag{5.1}$$

The identity $\zeta(s)^{-2} \cdot \zeta(s)^2 \equiv 1$ now leads to

$$\sum_{e|n} \beta\left(\frac{n}{e}\right) \cdot d(e) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \geq 2 \end{cases}. \tag{5.2}$$

It is now a simple matter (using formal Dirichlet series, or using (5.2) and complete induction over $N$) to prove that, if $a(N)$ and $b(N)$ are two complex-valued functions on $\mathbb{Z}^+$, then

$$\begin{aligned} a(N) &= \sum_{e|N} d\left(\frac{N}{e}\right) \cdot b(e), \quad \forall N \in \mathbb{Z}^+ \\ \iff \quad b(N) &= \sum_{e|N} \beta\left(\frac{N}{e}\right) \cdot a(e), \quad \forall N \in \mathbb{Z}^+. \end{aligned} \tag{5.3}$$

We notice some identities involving the function $\beta$. From $\zeta(s)\cdot\zeta(s)^{-2} \equiv \zeta(s)^{-1}$ we get

$$\sum_{e|n} \beta(e) = \mu(n), \qquad \forall n \in \mathbb{Z}^+. \tag{5.4}$$

We also have a relation involving the divisor sum function, $\sigma(n) := \sum_{e|n} e$ and the Euler phi-function, $\phi(n) = n\prod_{p|n}(1-1/p)$ :

$$\sum_{e|n} \beta\left(\frac{n}{e}\right)\sigma(e) = \phi(n), \qquad \forall n \in \mathbb{Z}^+. \tag{5.5}$$

This can be proved using the two formulas $\sum_{n=1}^{\infty} \phi(n)n^{-s} = \zeta(s-1)/\zeta(s)$ and $\sum_{n=1}^{\infty} \sigma(n)n^{-s} = \zeta(s)\zeta(s-1)$, true for Re $s > 2$ [T2, (1.2.12),(1.3.1)].

For $n \geq 2$ *square free*, and $p$ a prime dividing $n$, we let

$$S_p(n) := \sum_{\substack{e|n \\ (p|e)}} \beta\left(\frac{n}{e}\right) d(e), \tag{5.6}$$

where the sum is taken over those divisors of $n$ that are divisible by $p$. Substituting $f = e/p$, and using $d(e) = 2d(f)$ and (5.2), we find

$$S_p(n) = \begin{cases} 2 & \text{if } n = p \\ 0 & \text{otherwise} \end{cases}. \tag{5.7}$$

Using this, we can evaluate the following sum, for $n \geq 2$ square free:

$$\begin{aligned} \sum_{e|n} \beta\left(\frac{n}{e}\right) d(e)\cdot \ln e &= \sum_{e|n} \beta\left(\frac{n}{e}\right) d(e)\cdot \sum_{p|e} \ln p \\ &= \sum_{p|n} S_p(n)\cdot \ln p = \begin{cases} 2\ln n & \text{if } n \text{ is a prime} \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{5.8}$$

Now take $N \in \mathbb{Z}^+$ arbitrary, and $\lambda > 0$. Then by Theorem 4.6(a,c), $A(N,\lambda)$ decomposes as a direct sum

$$A(N,\lambda) = \bigoplus_{lM|N} \delta_l A_{\text{new}}(M,\lambda),$$

where $\langle l, M\rangle$ runs through all pairs of positive integers such that $lM \mid N$.

Now take $n \in \mathbb{Z}^+$ such that $(n,N) = 1$. By the remark following (4.1), the restriction of the Hecke operator ${}^N T_n$ to $A(M,\lambda)$ (where $M \mid N$) is simply ${}^M T_n$, and we will write $T_n$ instead of ${}^N T_n$.

By (4.14) and Lemma 4.4(e,g), $T_n$ preserves all spaces $\delta_l A_{\text{new}}(M,\lambda)$ occurring in the above direct sum, and $\text{Tr }(T_n)_{|\delta_l A_{\text{new}}(M,\lambda)} = \text{Tr }(T_n)_{|A_{\text{new}}(M,\lambda)}$. Hence we get

$$\text{Tr }(T_n)_{|A(N,\lambda)} = \sum_{M|N} d\left(\frac{N}{M}\right) \cdot \text{Tr }(T_n)_{|A_{\text{new}}(M,\lambda)}. \tag{5.9}$$

Keep $n \in \mathbb{Z}^+$ fixed. We invert (5.9) by using (5.3), where we can clearly replace "$\mathbb{Z}^+$" with "$\{N \in \mathbb{Z}^+ \mid (N,n) = 1\}$". We get, for each $N \in \mathbb{Z}^+$ relatively prime to $n$ :

$$\text{Tr }(T_n)_{|A_{\text{new}}(N,\lambda)} = \sum_{M|N} \beta(\frac{N}{M}) \cdot \text{Tr }(T_n)_{|A(M,\lambda)}. \tag{5.10}$$

For $n = 1$, (5.9) and (5.10) correspond to [AL, (6.6),(6.7)].

Notice that (5.10) is *not* true for $\lambda = 0$. From Definitions 1.4, 1.5 it is clear that $A_{\text{new}}(N,0) = \{0\}$ for all $N \geq 2$. On the other hand, we have by (2.8), (2.9):

$$\text{Tr }(T_n)_{|A(N,0)} = \frac{1}{\sqrt{n}} \sum_{d|n} d \qquad \text{when } (N,n) = 1.$$

Hence for $\lambda = 0$ we see (via (5.4)) that the right hand side of (5.10) equals $\mu(N) n^{-1/2} \sum_{d|n} d$.

We now take any function $h(r)$ satisfying (2.1), multiply (5.10) (adjusted when $\lambda = 0$) by $h(r_\lambda)$, and add over all eigenvalues $\lambda \geq 0$. We get:

$$\begin{aligned} & h\left(\frac{-i}{2}\right) \cdot \frac{\mu(N)}{\sqrt{n}} \sum_{d|n} d + \sum_{\lambda > 0} h(r_\lambda) \cdot \text{Tr }(T_n)_{|A_{\text{new}}(N,\lambda)} \\ & = \sum_{M|N} \beta(\frac{N}{M}) \cdot \left( \sum_{\lambda \geq 0} h(r_\lambda) \cdot \text{Tr }(T_n)_{|A(M,\lambda)} \right). \end{aligned} \tag{5.11}$$

This holds for any $N, n \in \mathbb{Z}^+$ such that $(N,n) = 1$.

Now add the assumption, that $N$ be *square free* and $\geq 2$. Then each divisor $M$ of $N$ is also square free, and we can evaluate the right hand side of (5.11) by using the trace formula on p. 19, with the explicit form from Theorem 3.9 for the third and fourth line.

For the identity term on p. 19, notice that for $M$ square free, $M \cdot \prod_{p|M}(1 + p^{-1})$ equals $\sigma(M)$, the divisor sum function. Thus the contribution to (5.11) from the identity term can be evaluated via (5.5), giving:

$$\left[\underline{\text{if } \sqrt{n} \in \mathbb{Z}:} \quad \frac{\prod_{p|N}(p-1)}{12\sqrt{n}} \int_{-\infty}^{\infty} rh(r)\tanh(\pi r)\,dr\right].$$

To evaluate the contribution from the expression in Theorem 3.9, we need to compute, for arbitrary $d$ as in (3.1) and $f \in \mathbb{Z}^+$ :

$$\begin{aligned}
&\sum_{M|N} \beta(\frac{N}{M}) \cdot \prod_{p|M} \begin{Bmatrix} 2 & \text{if } p \mid f \\ 1+(d/p) & \text{if } p \nmid f \end{Bmatrix} \\
&= \sum_{M|N} \Bigg(\prod_{p|\frac{N}{M}} (-2)\Bigg) \cdot \prod_{p|M} \begin{Bmatrix} 2 & \text{if } p \mid f \\ 1+(d/p) & \text{if } p \nmid f \end{Bmatrix} \\
&= \sum_{M|N} \prod_{p|N} \begin{Bmatrix} 2 & \text{if } p \mid M,\ p \mid f \\ 1+(d/p) & \text{if } p \mid M,\ p \nmid f \\ -2 & \text{if } p \nmid M \end{Bmatrix} \\
&= \prod_{p|N} \left( \begin{Bmatrix} 2 & \text{if } p \mid f \\ 1+(d/p) & \text{if } p \nmid f \end{Bmatrix} - 2 \right) \\
&= \prod_{p|N} \begin{Bmatrix} 0 & \text{if } p \mid f \\ (d/p)-1 & \text{if } p \nmid f \end{Bmatrix}
\end{aligned}$$

Using this we see that the contribution from the expression in Theorem 3.9 is:

$$\sum_{\substack{t\in\mathbb{Z} \\ \sqrt{t^2-4n}\notin\mathbb{Q}}} \left[ \sum_{\substack{f|l \\ (f,N)=1}} \mathbf{h}(\mathfrak{r}[f]) \cdot \left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right] \cdot \prod_{p|N} \left( \left(\frac{d}{p}\right) - 1 \right) \right] \cdot A(t,n).$$

The contribution from the remaining terms on p. 19, *except* for the $[\ln Y(|a-d|, N)]$-term and the last line, is easy to evaluate, using (5.2) and (5.8). The result is 0 if $N$ is composite, and if $N = p$ is a prime we get:

$$\left[\text{if } \sqrt{n} \in \mathbb{Z}: \quad -2(\ln p) g(0)\right] - \ln p \cdot \sum_{\substack{ad=n \\ a\neq d}} g(\ln \frac{a}{d}).$$

The $[\ln Y(|a-d|,N)]$-term on p. 19 gives the following contribution to (5.11):

$$\begin{aligned}
&\sum_{M|N}\beta(\frac{N}{M})d(M)\sum_{\substack{ad=n\\a\neq d}}\frac{-\ln Y(|a-d|,M)}{2|a-d|}\cdot g(\ln\frac{a}{d})\\
&=-\sum_{\substack{ad=n\\a\neq d}}\frac{g(\ln\frac{a}{d})}{2|a-d|}\sum_{M|N}\beta(\frac{N}{M})d(M)\cdot\sum_{k\bmod |a-d|}\ln\left(M,\frac{|a-d|}{(k,|a-d|)}\right)\\
&=-\sum_{\substack{ad=n\\a\neq d}}\frac{g(\ln\frac{a}{d})}{2|a-d|}\sum_{M|N}\beta(\frac{N}{M})d(M)\cdot\sum_{k\bmod |a-d|}\sum_{p}{}'\ln p,
\end{aligned}$$

where the $\sum_p'$-sum is taken over all primes $p$ that divides both $M$ and $|a-d|/(k,|a-d|)$. Using the notation from (5.6), we can write this as

$$-\sum_{\substack{ad=n\\a\neq d}}\frac{g(\ln\frac{a}{d})}{2|a-d|}\sum_{k\bmod |a-d|}\sum_{p}{}'S_p(N)\ln p,$$

where the $\sum_p'$-sum is now taken over all primes $p$ that divides both $N$ and $|a-d|/(k,|a-d|)$. By (5.7), this is 0 if $N$ is composite, and if $N=p$ a prime we get:

$$\begin{aligned}
&-\sum_{\substack{ad=n\\a\neq d}}\frac{g(\ln\frac{a}{d})}{2|a-d|}\cdot\left|\left\{k\bmod |a-d|\quad\middle|\quad p\mid\frac{|a-d|}{(k,|a-d|)}\right\}\right|\cdot 2\ln p\\
&=-\sum_{\substack{ad=n\\a\neq d}}\frac{g(\ln\frac{a}{d})}{2|a-d|}\cdot|a-d|\cdot\left(1-\frac{1}{(p^\infty,|a-d|)}\right)\cdot 2\ln p\\
&=-\ln p\sum_{\substack{ad=n\\a\neq d}}\left(1-\frac{1}{(p^\infty,|a-d|)}\right)g(\ln\frac{a}{d}).
\end{aligned}$$

Finally, the last line in the formula on p. 19 gives the following contribution to (5.11):

$$\begin{aligned}
&\sum_{M|N}\beta(\frac{N}{M})\cdot 2d(M)\sum_{m=2}^{\infty}\frac{\Lambda(m)}{m}\sum_{ad=n}g\left(\ln(\frac{a}{d})-2\ln m\right)\\
&-\sum_{M|N}\beta(\frac{N}{M})\cdot d(M)\sum_{p|M}\sum_{r=1}^{\infty}\frac{\ln p}{p^r}\sum_{ad=n}g\left(\ln(\frac{a}{d})-2\ln p^r\right).
\end{aligned}$$

Here the first line is 0, by (5.2). The second line is, again using the notation from (5.6):

$$-\sum_{p|N} S_p(N) \sum_{r=1}^{\infty} \frac{\ln p}{p^r} \sum_{ad=n} g\left(\ln(\frac{a}{d}) - 2\ln p^r\right).$$

By (5.7), this is 0 if $N$ is composite, and if $N = p$ a prime we get:

$$-2\sum_{r=1}^{\infty} \frac{\ln p}{p^r} \sum_{ad=n} g\left(\ln(\frac{a}{d}) - 2\ln p^r\right).$$

We now collect our results. We have proved that for any function $h(r)$ satisfying (2.1), and any $N, n \in \mathbb{Z}^+$ such that $(N, n) = 1$, $N$ square free and $N \geq 2$, the following formula holds:

(5.12)

$$h\left(\frac{-i}{2}\right) \cdot \frac{\mu(N)}{\sqrt{n}} \sum_{d|n} d + \sum_{\lambda>0} h(r_\lambda) \cdot \text{Tr } (T_n)_{|A_{\text{new}}(N,\lambda)}$$
$$= \left[\underline{\text{if } \sqrt{n} \in \mathbb{Z}:} \quad \frac{\prod_{p|N}(p-1)}{12\sqrt{n}} \int_{-\infty}^{\infty} rh(r)\tanh(\pi r)\, dr\right]$$
$$+ \sum_{\substack{t \in \mathbb{Z} \\ \sqrt{t^2-4n} \notin \mathbb{Q}}} \left[\sum_{\substack{f|l \\ (f,N)=1}} \mathbf{h}(\mathfrak{r}[f]) \cdot \left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right] \cdot \prod_{p|N} \left(\left(\frac{d}{p}\right) - 1\right)\right] \cdot A(t,n)$$
$$- \Bigg[\underline{\text{if } N = p, \text{ a prime}:} \quad 2\ln p \sum_{ad=n} g(\ln\frac{a}{d}) - \ln p \sum_{\substack{ad=n \\ a \neq d}} \frac{g(\ln\frac{a}{d})}{(p^\infty, |a-d|)}$$
$$+ 2\sum_{r=1}^{\infty} \frac{\ln p}{p^r} \sum_{ad=n} g\left(\ln(\frac{a}{d}) - 2\ln p^r\right)\Bigg].$$

Here, in line three, $d, l$ are "implicit variables" depending on $t$ as in (3.3), i.e. $t^2 - 4n = l^2 d$, $l \in \mathbb{Z}^+$, and $d$ is a fundamental discriminant (3.1). Also, $\mathfrak{r}[f] \subset \mathbb{Q}(\sqrt{d})$ is as in (3.4). For $\mathbf{h}(\mathfrak{r}[f])$, cf. (3.12), (3.13).

*Remark* 5.1. If [BV, Thm. 2] generalizes in some way from $\Gamma(2) \cong \Gamma_0(4)$ to Hecke congruence subgroups of arbitrary level, then we could probably give an alternative proof of (5.12) by using the trace formula for modular correspondences on Fuchsian groups *with characters* (to appear in Part II of [S5]).

When $N$ equals the reduced discriminant $d_B$ of an indefinite division quaternion algebra, formula (5.12) agrees perfectly with Theorem 3.11! This gives:

**Theorem 5.2.** *Let $B$ be an indefinite division quaternion algebra over $\mathbb{Q}$, and let $R$ be a maximal order in $B$. Let $d_B > 1$ be the reduced discriminant of $B$. Take $n \in \mathbb{Z}^+$ such that $(n, d_B) = 1$. Write $\widetilde{T}_n$ for the Hecke operator* (1.11) *on the cocompact group $\Gamma_R$. Let $T_n$ be the Hecke operator* (2.9) *on $\Gamma_0(d_B)$. Then, for any function $h(r)$ as in* (2.1)*:*

$$h\left(\frac{-i}{2}\right) \cdot \frac{1}{\sqrt{n}} \sum_{d|n} d + \sum_{\lambda>0} h(r_\lambda) \cdot Tr\ (T_n)_{|A_{new}(d_B,\lambda)} = \sum_{\lambda \geq 0} h(r_\lambda) \cdot Tr\ (\widetilde{T}_n)_{|A(\Gamma_R \backslash \mathcal{H},\lambda)}. \tag{5.13}$$

*Proof.* Simply combine (5.12) and Theorem 3.11. Remember that $d_B$ is square free, $> 1$, and has an *even* number of prime factors [V, p. 74, Thm. 3.1]. $\square$

**Corollary 5.3.** *Take $n \in \mathbb{Z}^+$ such that $(n, d_B) = 1$. Then for each $\lambda > 0$ we have*

$$Tr\ (T_n)_{|A_{new}(d_B,\lambda)} = Tr\ (\widetilde{T}_n)_{|A(\Gamma_R \backslash \mathcal{H},\lambda)}. \tag{5.14}$$

*Also,*

$$Tr\ (\widetilde{T}_n)_{|A(\Gamma_R \backslash \mathcal{H},0)} = \frac{1}{\sqrt{n}} \sum_{d|n} d. \tag{5.15}$$

*Proof.* Fix some $r_0 \in i[-\frac{1}{2}, 0] \cup [0, \infty)$, and let $\lambda_0 = \frac{1}{4} + r_0^2$. Let $\Delta$ be a positive number such that $|\lambda - \lambda_0| > \Delta$ for all discrete eigenvalues $\lambda \neq \lambda_0$ on $\Gamma_0(d_B)$ and $\Gamma_R$. Define, for $C > 0$:

$$h_C(r) = \exp\left[-C(r^2 - r_0^2)^2\right]. \tag{5.16}$$

Clearly, $h_C(r)$ satisfies (2.1), and $0 < h_C(r) \leq 1$ for $r \in i\mathbb{R} \cup \mathbb{R}$. Notice $h_C(r_0) = 1$, for all $C$. Also,

$$|\lambda - \lambda_0| > \Delta \text{ and } C > 2 \quad \Longrightarrow \quad h_C(r_\lambda) < e^{-C\Delta^2/2} \cdot h_1(r_\lambda). \tag{5.17}$$

By (5.10), [H4, p. 315(ii)], [H3, Ch. 5, Prop. 2.10(b)], [S5, Lemma 2.1(c)], both sums in (5.13) are absolutely convergent for any $h \equiv h_C$ (cf. [S5, proof of Thm. 8.3]). We use this fact for $h \equiv h_1$. Hence, by (5.17), if we take $C \to \infty$ in (5.13), we obtain, for $\lambda_0 > 0$ :

$$\text{Tr}\ (T_n)_{|A_{\text{new}}(d_B,\lambda_0)} = \text{Tr}\ (\widetilde{T}_n)_{|A(\Gamma_R \backslash \mathcal{H},\lambda_0)}.$$

Hence (5.14) is proved. Taking $\lambda_0 = 0$ (i.e. $r_0 = -\frac{i}{2}$) we obtain (5.15). $\square$

*Remark* 5.4. The numbers $\widetilde{\tau}_0(n) = \text{Tr } (\widetilde{T}_n)_{|A(\Gamma_R\backslash\mathcal{H},0)}$ can of course be determined for *all* $n \in \mathbb{Z}^+$. From (1.11) and [S2, p. 283(7)] we see that:

$$\widetilde{\tau}_0(n) = \frac{1}{\sqrt{n}} \sum_{\substack{d|n \\ (d,d_B)=1}} d.$$

*Remark* 5.5. We end this section by giving an alternative proof of Corollary 5.3, which shows that by using some simple functional analysis, a significant amount of explicit computations in this section and in Section 2 could have been avoided.

To get started, we claim that the general trace formula (2.2) can be written as:

$$\begin{aligned} &\sum_{\lambda\geq 0} h(r_\lambda) \cdot \text{Tr } \mathcal{M}_{|A(\lambda)} \\ &= \frac{1}{2} \sum_{\substack{T\in H^{-1}//\Gamma, \\ \text{hyperbolic,} \\ \text{doesn't fix cusp}}} \frac{\ln \mathcal{N}(T_0)}{\mathcal{N}(T)^{1/2} - \mathcal{N}(T)^{-1/2}} \cdot g(\ln \mathcal{N}(T)) \\ &+ \int_{-\infty}^{\infty} F(r)h(r)\, dr + C \cdot h(0), \end{aligned} \tag{5.18}$$

where $C$ is some constant, and $F(r)$ is some smooth function. $C$ and $F(r)$ depend only on $\Gamma$ and $\mathcal{M}$, and we have nice majorized convergence $\int_{-\infty}^{\infty} |F(r)h(r)|\, dr < \infty$ for any function $h(r)$ satisfying (2.1).

When proving this claim, most terms in (2.2) can be expressed in the desired form immediately (recall that $g(u)$ is the Fourier transform of $h(r)$). Notice that there are only finitely many elliptic terms in (2.2) (see [S5, Thm. 3.6], [H3, Ch. 5, Prop. 3.2]), so these can all be accounted for in the $\int_{-\infty}^{\infty} F(r)h(r)\, dr$–term. The integral in line 7 of (2.2) is rewritten as follows (with $\gamma = \gamma_{ij} \in (0,\infty) - \{1\}$):

$$\begin{aligned} &\int_{|\ln\gamma|}^{\infty} g(u) \cdot \frac{e^{u/2} + e^{-u/2}}{e^{u/2} - e^{-u/2} + |\gamma^{1/2} - \gamma^{-1/2}|}\, du \\ &= \int_0^{\infty} g(u)\, du - \int_0^{|\ln\gamma|} g(u)\, du + \int_{|\ln\gamma|}^{\infty} g(u) \frac{2e^{-u/2} - |\gamma^{1/2} - \gamma^{-1/2}|}{e^{u/2} - e^{-u/2} + |\gamma^{1/2} - \gamma^{-1/2}|}\, du \end{aligned}$$

(5.19)

$$= \frac{1}{2}h(0) + \int_{-\infty}^{\infty} h(r)\left(\frac{1}{2\pi}\int_{|\ln\gamma|}^{\infty} e^{-iru}\frac{2e^{-u/2} - |\gamma^{1/2} - \gamma^{-1/2}|}{e^{u/2} - e^{-u/2} + |\gamma^{1/2} - \gamma^{-1/2}|}\,du \right.$$
$$\left. - \frac{1}{2\pi}\int_0^{|\ln\gamma|} e^{-iru}\,du\right) dr.$$

Notice that the expression in the big parenthesis is certainly a bounded function of $r$, for fixed $\gamma \in (0,\infty) - \{1\}$. Finally for line 8 of (2.2), majorized convergence was shown in [S5, Thm 8.3].

From (2.2) and (5.19) we see that the constant $C$ in (5.18) is in fact:

$$\begin{aligned} C =& \frac{1}{4}\sum_{i,j}^{(p)} 1 + \frac{1}{4}\sum_{i,j}^{(h)} \sqrt{\gamma_{ij}} - \frac{1}{4}\sum_{j,l}\varphi_{jl}(\frac{1}{2})\xi_{lj}(\frac{1}{2}) \\ =& \frac{1}{4}\sum_{j=1}^{\kappa}\sum_{\substack{i=1\\(k(i,j)=j)}}^{d}\sqrt{\gamma_{ij}} - \frac{1}{4}\sum_{j=1}^{\kappa}\sum_{l=1}^{\kappa}\varphi_{jl}(\frac{1}{2})\xi_{lj}(\frac{1}{2}) \\ =& \frac{1}{4}\sum_{j=1}^{\kappa}\sum_{l=1}^{\kappa}\left(\delta_{jl} - \varphi_{jl}(\frac{1}{2})\right)\xi_{lj}(\frac{1}{2}). \qquad \{\text{By (2.6).}\} \end{aligned}$$

In the case of the Hecke operator $T_n$ on $\Gamma = \Gamma_0(N)$ ($N$ square free, and $(n,N) = 1$), we can evaluate this as in Section 2. We get, after dividing by $\sqrt{n}$ from (2.9):

$$C = \frac{1}{2}d(N)\sum_{ad=n} 1.$$

Using this, (5.11), (5.2), and our computations for the *hyperbolic* terms from Section 3 and p. 50, we arrive at the following version of formula (5.12) (for $N, n \in \mathbb{Z}^+$, $(N,n) = 1$, $N$ square free and $N \geq 2$):

$$h\left(\frac{-i}{2}\right)\cdot\frac{\mu(N)}{\sqrt{n}}\sum_{d|n} d + \sum_{\lambda>0} h(r_\lambda)\cdot \mathrm{Tr}\ (T_n)_{|A_{\mathrm{new}}(N,\lambda)}$$
$$= \sum_{\substack{t\in\mathbb{Z}\\ t^2-4n>0\\ \sqrt{t^2-4n}\notin\mathbb{Q}}}\left[\sum_{\substack{f|l\\(f,N)=1}} \mathbf{h}(\mathfrak{r}[f])\cdot\left[\mathfrak{r}[1]^1 : \mathfrak{r}[f]^1\right]\cdot\prod_{p|N}\left(\left(\frac{d}{p}\right) - 1\right)\right]\cdot A(t,n)$$
$$+ \int_{-\infty}^{\infty} F(r)h(r)\,dr.$$

Here $F(r)$ is some smooth function that only depends on $N, n$, and the last integral is nicely majorized for any $h(r)$ as in (2.1).

Combining this with Theorem 3.11 (where we only use the explicit form for the *hyperbolic* terms), we get, for any $n \in \mathbb{Z}^+$ relatively prime to $d_B$, the reduced discriminant of $B$ :

$$h\left(\frac{-i}{2}\right) \cdot \frac{1}{\sqrt{n}} \sum_{d|n} d + \sum_{\lambda>0} h(r_\lambda) \cdot \text{Tr } (T_n)_{|A_{\text{new}}(d_B,\lambda)} - \sum_{\lambda \geq 0} h(r_\lambda) \cdot \text{Tr } (\widetilde{T}_n)_{|A(\Gamma_R \backslash \mathcal{H},\lambda)} = \int_{-\infty}^{\infty} F(r) h(r)\, dr. \tag{5.20}$$

(With a *new* function $F(r)$.)

We can now repeat the argument from the proof of Corollary 5.3. The only new thing to observe is that for fixed $r_0 \in i[-\frac{1}{2}, 0] \cup [0, \infty)$, and $h_C(r)$ defined as in (5.16), we have

$$\int_{-\infty}^{\infty} |F(r) h_C(r)|\, dr \to 0 \quad \text{as } C \to \infty. \tag{5.21}$$

To see this, notice that for any $\Delta > 0$, the inequality (5.17) implies, for $C > 2$:

$$\int_{-\infty}^{\infty} |F(r) h_C(r)|\, dr \leq \int_{|r^2 - r_0^2| \leq \Delta} |F(r) h_C(r)|\, dr + e^{-C\Delta^2/2} \int_{-\infty}^{\infty} |F(r) h_1(r)|\, dr.$$

(The first integral in the right hand side may be empty, if $r_0 \notin \mathbb{R}$.) Here the last integral is known to be absolutely convergent. Since $h_C(r) \leq 1$ for all $r \in \mathbb{R}$, we obtain

$$\limsup_{C \to \infty} \int_{-\infty}^{\infty} |F(r) h_C(r)|\, dr \leq \int_{|r^2 - r_0^2| \leq \Delta} |F(r)|\, dr.$$

But the right hand side can be made arbitrarily small, by taking $\Delta$ small enough. Hence (5.21) is proved. Hence, arguing as in the proof of Corollary 5.3, we see that (5.20) implies (5.14), (5.15).

## 6. Proof of the main theorems

In this section we prove our main theorems, i.e. Theorem 1.6, Theorem 1.8 and Theorem 1.9. These theorems will follow fairly directly from Corollary 5.3 in last section, together with the multiplicativity of the Hecke operators.

Our notation will be the same as in Theorem 5.2, i.e. $B$ is an indefinite division quaternion algebra over $\mathbb{Q}$, and $R$ is a maximal order in

$B$. The reduced discriminant of $B$ is $d_B > 1$. We write $\widetilde{T}_n$ ($n \in \mathbb{Z}^+$) for the Hecke operator (1.11) on the cocompact group $\Gamma_R$, and ${}^{d_B}T_n$ (or simply $T_n$, if $(n, d_B) = 1$) for the Hecke operator (2.9) on $\Gamma_0(d_B)$.

We will give the proof in a series of lemmas. Concerning the $\Theta$-map from Section 1 (see Theorem 1.3) we have:

**Lemma 6.1.** *For any $\varphi \in A(\Gamma_R \setminus \mathcal{H}, \lambda)$ ($\lambda > 0$):*

$$\Theta(\mathcal{P}\varphi) = V\left[\Theta(\varphi)\right], \quad \textit{and} \quad \Theta(\widetilde{T}_n\varphi) = {}^{d_B}T_n\left[\Theta(\varphi)\right], \quad \forall n \in \mathbb{Z}^+.$$

*Proof.* The first assertion follows at once from Theorem 1.3, (1.13), (1.14) and Lemma 4.4(d). For the second assertion, we take the proof from [BJ1, pp. 65-66], somewhat simplified. Define $c(m)$ as the Fourier coefficients of $\Theta(\varphi)$, as in Theorem 1.3. Applying Theorem 1.3 to $\widetilde{T}_n\varphi$, we get

$$\Theta(\widetilde{T}_n\varphi)(\tau) = \sum_{m \neq 0} c'(m)\sqrt{v}K_{ir}(2\pi|m|v)e^{2\pi imu} \qquad (\tau = u + iv),$$

where $c'(m) = 4 \cdot \begin{cases} \left(\widetilde{T}_m(\widetilde{T}_n\varphi)\right)(z_0) & \text{if } m \geq 1 \\ \left(\mathcal{P}\widetilde{T}_{|m|}(\widetilde{T}_n\varphi)\right)(z_0) & \text{if } m \leq -1 \end{cases}.$

From Lemma 1.7(c,d) one can prove, just as in Remark 4.3:

$$\widetilde{T}_m \circ \widetilde{T}_n = \sum_{\substack{d|(m,n) \\ (d,d_B)=1}} \widetilde{T}_{mn/d^2}.$$

Using this we obtain

$$c'(m) = \sum_{\substack{d|(m,n) \\ (d,d_B)=1 \\ d\in\mathbb{Z}^+}} c(\frac{mn}{d^2}), \quad \text{for all } m \neq 0.$$

Hence by Lemma 4.1, $\Theta(\widetilde{T}_n\varphi) = {}^{d_B}T_n\left[\Theta(\varphi)\right]$, as asserted. □

We now start deriving consequences of Corollary 5.3.

For $n = 1$, Corollary 5.3 gives

$$\dim A_{\text{new}}(d_B, \lambda) = \dim A(\Gamma_R \setminus \mathcal{H}, \lambda), \qquad \forall \lambda > 0.$$

Now *fix* some $\lambda > 0$ for which this dimension is $D > 0$. By Theorem 4.6(a), the newforms $f_1, ..., f_D$ in $A_{\text{new}}(d_B, \lambda)$ form an orthogonal basis in $A_{\text{new}}(d_B, \lambda)$. By Theorem 4.6(b) and Definition 4.5(i) we have

(6.1) $$Vf_j = \epsilon_j \cdot f_j, \qquad {}^{d_B}T_n f_j = \tau_j(n) \cdot f_j, \quad \forall n \in \mathbb{Z}^+,$$

for some $\epsilon_j \in \{-1, 1\}$, $\tau_j(n) \in \mathbb{R}$. Of course, $\tau_j(1) = 1$.

We also choose a Hecke basis in $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ as in (1.19):

$$\begin{aligned} &\{\widetilde{f}_1, ..., \widetilde{f}_D\} \text{ is an orthonormal basis in } A(\Gamma_R \setminus \mathcal{H}, \lambda); \\ &\mathcal{P}\widetilde{f}_j = \widetilde{\epsilon}_j \cdot \widetilde{f}_j, \qquad \widetilde{T}_n \widetilde{f}_j = \widetilde{\tau}_j(n) \cdot \widetilde{f}_j, \quad \forall n \in \mathbb{Z}^+. \end{aligned} \tag{6.2}$$

We know that $\widetilde{\epsilon}_j \in \{-1, 1\}$, $\widetilde{\tau}_j(n) \in \mathbb{R}$ and $\widetilde{\tau}_j(1) = 1$.

For each $n \in \mathbb{Z}^+$ relatively prime to $d_B$ we have, by Corollary 5.3:

$$\sum_{j=1}^{D} \tau_j(n) = \sum_{j=1}^{D} \widetilde{\tau}_j(n). \tag{6.3}$$

Using multiplicativity, this gives:

**Lemma 6.2.** *Let $p_1, ..., p_M$ be distinct primes not dividing $d_B$, and take $m_1, ..., m_M \in \{0, 1, 2, 3, ...\}$. Then*

$$\sum_{j=1}^{D} \prod_{i=1}^{M} \tau_j(p_i)^{m_i} = \sum_{j=1}^{D} \prod_{i=1}^{M} \widetilde{\tau}_j(p_i)^{m_i}.$$

*Proof.* From Lemma 4.2 we know that

$$\tau_j(n_1) \cdot \tau_j(n_2) = \tau_j(n_1 n_2) \qquad \text{when } (n_1, n_2) = 1, \tag{6.4}$$

$$\tau_j(p) \cdot \tau_j(p^e) = \tau_j(p^{e+1}) + \tau_j(p^{e-1}) \qquad \text{for } e \geq 1, p \nmid d_B. \tag{6.5}$$

From (6.5) we can compute, for primes $p$ not dividing $d_B$ :

$$\begin{aligned} &\tau_j(p)^2 = \tau_j(p^2) + 1; \\ &\tau_j(p)^3 = \big(\tau_j(p^2) + 1\big) \cdot \tau_j(p) = \tau_j(p^3) + 2\tau_j(p); \\ &\tau_j(p)^4 = \big(\tau_j(p^3) + 2\tau_j(p)\big) \cdot \tau_j(p) = \tau_j(p^4) + 3\tau_j(p^2) + 2; \\ &\qquad . \qquad . \qquad . \qquad . \end{aligned}$$

Also, $\tau_j(p)^0 = 1 = \tau_j(p^0)$. In this way, we see that for any $m \in \{0, 1, 2, 3, ...\}$ :

$$\tau_j(p)^m = \sum_{k=0}^{m} c_{mk} \cdot \tau_j(p^k),$$

where each $c_{mk}$ is an integer that only depends on $m$ and $k$. Using (6.4) we then get:

$$\begin{aligned}\prod_{i=1}^{M}\tau_j(p_i)^{m_i} &= \prod_{i=1}^{M}\left(\sum_{k=0}^{m_i} c_{m_ik}\cdot\tau_j(p_i^k)\right)\\ &= \sum_{k_1=0}^{m_1}\dots\sum_{k_M=0}^{m_M}\prod_{i=1}^{M}\left(c_{m_ik_i}\cdot\tau_j(p_i^{k_i})\right)\\ &= \sum_{k_1=0}^{m_1}\dots\sum_{k_M=0}^{m_M}\left(\prod_{i=1}^{M}c_{m_ik_i}\right)\cdot\tau_j(p_1^{k_1}\cdot\ldots\cdot p_M^{k_M}).\end{aligned}$$

But because of Lemma 1.7(c,d), the $\widetilde{\tau}_j$'s satisfy exactly the same multiplicativity relations as the $\tau_j$'s, and hence the above equalities (with *the same* integers $c_{mk}$) remain true if we replace $\tau_j$ with $\widetilde{\tau}_j$ throughout. Now our assertion follows from (6.3). $\square$

**Lemma 6.3.** *There is a permutation $\sigma$ of $\{1,...,D\}$ such that $\tau_{\sigma(j)}(p)=\widetilde{\tau}_j(p)$ for each $j\in\{1,...,D\}$ and each prime $p$ not dividing $d_B$.*

*Proof.* Let $p_1,p_2,p_3,...$ be an enumeration of *all* prime numbers not dividing $d_B$. For each $M\in\mathbb{Z}^+$, let $S_M$ be the set of permutations $\sigma$ of $\{1,...,D\}$ such that $\tau_{\sigma(j)}(p_k)=\widetilde{\tau}_j(p_k)$ for each $j\in\{1,...,D\}$ and each $k\in\{1,...,M\}$. It is enough to prove that each set $S_M$ is nonempty. For if this is the case, we can take some $\sigma_M\in S_M$ for each $M$, and since there only exist finitely many permutations of $\{1,2,...,D\}$, some permutation $\sigma$ must occur infinitely many times in the sequence $\sigma_1,\sigma_2,...$ But then, since $S_1\supset S_2\supset S_3\supset...$, this $\sigma$ will in fact belong to *all* sets $S_M$, and hence $\tau_{\sigma(j)}(p)=\widetilde{\tau}_j(p)$ for all $j\in\{1,...,D\}$ and all primes $p$ not dividing $d_B$.

Now fix $M\in\mathbb{Z}^+$. To show that $S_M$ is nonempty, it is in fact enough to show that

$$\begin{aligned}&\left|\{j\in\{1,2,...,D\}\mid(\tau_j(p_1),...,\tau_j(p_M))=(a_1,...,a_M)\}\right|\\ &\quad=\left|\{j\in\{1,2,...,D\}\mid(\widetilde{\tau}_j(p_1),...,\widetilde{\tau}_j(p_M))=(a_1,...,a_M)\}\right|,\end{aligned}\tag{6.6}$$

for each vector $(a_1,...,a_M)\in\mathbb{R}^M$. For assume that (6.6) holds. We can then construct a permutation $\sigma\in S_M$ by defining $\sigma(1),\sigma(2),...$, one at a time, in the following way. Take $k\in\{1,2,...,D\}$. If $k\geq 2$, we assume that $\sigma(1),...,\sigma(k-1)$ have been assigned pairwise different values in $\{1,2,...,D\}$, so that $(\widetilde{\tau}_j(p_1),...,\widetilde{\tau}_j(p_M))=(\tau_{\sigma(j)}(p_1),...,\tau_{\sigma(j)}(p_M))$ for each $j=1,...,k-1$. We want to define $\sigma(k)$.

Taking $(a_1, ..., a_M) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))$ in (6.6) we get:

$$\begin{aligned}
&\Big|\{j \in \{1,2,...,D\} \mid (\tau_j(p_1), ..., \tau_j(p_M)) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))\}\Big| \\
&= \Big|\{j \in \{1,2,...,D\} \mid (\widetilde{\tau}_j(p_1), ..., \widetilde{\tau}_j(p_M)) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))\}\Big|, \\
&\geq 1 + \Big|\{j \in \{1,2,...,k-1\} \mid (\widetilde{\tau}_j(p_1), ..., \widetilde{\tau}_j(p_M)) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))\}\Big| \\
&= 1 + \Big|\{j \in \{1,2,...,k-1\} \mid (\tau_{\sigma(j)}(p_1), ..., \tau_{\sigma(j)}(p_M)) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))\}\Big|.
\end{aligned}$$

This shows that there is some $j \in \{1,2,...,D\}$, not occurring among the already assigned values $\sigma(1), ..., \sigma(k-1)$, such that $(\tau_j(p_1), ..., \tau_j(p_M)) = (\widetilde{\tau}_k(p_1), ..., \widetilde{\tau}_k(p_M))$. Then define $\sigma(k) := j$.

Repeating this for $k = 1, 2, ..., D$, we finally obtain a permutation $\sigma \in S_M$, showing that $S_M$ is nonempty.

Now, it only remains to prove (6.6). For this, we will use Lemma 6.2.

From Lemma 6.2 we see that

$$\sum_{j=1}^{D} P(\tau_j(p_1), ..., \tau_j(p_M)) = \sum_{j=1}^{D} P(\widetilde{\tau}_j(p_1), ..., \widetilde{\tau}_j(p_M)) \tag{6.7}$$

for any polynomial $P(X_1, X_2, ..., X_M)$ with real (or complex) coefficients. In this relation, we take $P(X_1, ..., X_M)$ to be the polynomial

$$P(X_1, ..., X_M) = \prod_{j=1}^{D} \prod_{\substack{1 \leq k \leq M \\ \tau_j(p_k) \neq a_k}} (X_k - \tau_j(p_k)) \cdot \prod_{j=1}^{D} \prod_{\substack{1 \leq k \leq M \\ \widetilde{\tau}_j(p_k) \neq a_k}} (X_k - \widetilde{\tau}_j(p_k)).$$

(As usual, an empty product is defined to be 1.)

We let $K = P(a_1, a_2, ..., a_M)$. Clearly, $K \neq 0$. On the other hand, for any $j \in \{1,2,...,D\}$ such that $(\tau_j(p_1), ..., \tau_j(p_M)) \neq (a_1, ..., a_M)$, there is some $k \in \{1,2,...,M\}$ such that $\tau_j(p_k) \neq a_k$. Then the factor $(X_k - \tau_j(p_k))$ occurs in $P(X_1, ..., X_M)$, and thus $P(\tau_j(p_1), ..., \tau_j(p_M)) = 0$. Hence, for our choice of $P(X_1, ..., X_M)$ we have:

$$\begin{aligned}
&\sum_{j=1}^{D} P(\tau_j(p_1), ..., \tau_j(p_M)) \\
&= \Big|\{j \in \{1,2,...,D\} \mid (\tau_j(p_1), ..., \tau_j(p_M)) = (a_1, ..., a_M)\}\Big| \cdot K.
\end{aligned}$$

Similarly:

$$\sum_{j=1}^{D} P(\widetilde{\tau}_j(p_1), ..., \widetilde{\tau}_j(p_M))$$
$$= \Big| \{ j \in \{1, 2, ..., D\} \mid (\widetilde{\tau}_j(p_1), ..., \widetilde{\tau}_j(p_M)) = (a_1, ..., a_M) \} \Big| \cdot K.$$

Together with (6.7) (and $K \neq 0$), this proves (6.6).

□

Because of this lemma, by simply reordering $f_1, ..., f_D$, we may assume $\tau_j(p) = \widetilde{\tau}_j(p)$ for each $j \in \{1, ..., D\}$ and each prime $p$ not dividing $d_B$. Now take some $j \in \{1, ..., D\}$, and look at $\Theta(\widetilde{f}_j)$. By Theorem 1.3 we have $\Theta(\widetilde{f}_j) \in A(d_B, \lambda)$. Also, by (6.2) and Lemma 6.1:

$$\begin{cases} V\left(\Theta(\widetilde{f}_j)\right) = \widetilde{\epsilon}_j \cdot \Theta(\widetilde{f}_j) \\ {}^{d_B}T_n\left(\Theta(\widetilde{f}_j)\right) = \widetilde{\tau}_j(n) \cdot \Theta(\widetilde{f}_j), \qquad \forall n \in \mathbb{Z}^+. \end{cases} \tag{6.8}$$

In particular, we have $T_p\left(\Theta(\widetilde{f}_j)\right) = \tau_j(p) \cdot \left(\Theta(\widetilde{f}_j)\right)$ for each prime $p$ not dividing $d_B$. We claim that this, together with our results in Section 4, implies:

$$\Theta(\widetilde{f}_j) \text{ is a constant multiple of the newform } f_j. \tag{6.9}$$

To see this[18]: Expand $\Theta(\widetilde{f}_j)$ in the basis of $A(d_B, \lambda)$ given by Theorem 4.6(c). By Lemma 4.4(g) and the remark following (4.1), each element in this basis is an eigenfunction to each $T_p$ ($p \nmid d_B$). The newform $f_j$ is an element in this basis, and by Theorem 4.6(d), this is the only basis element which has the same $T_p$-eigenvalue as $\Theta(\widetilde{f}_j)$ for each $p \nmid d_B$. Hence it is the only basis element which can occur nonvanishingly in our expansion of $\Theta(\widetilde{f}_j)$; i.e. (6.9) holds.

Since (6.9) holds for each $j \in \{1, ..., D\}$, we see that $\Theta$ maps $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ *into* $A_{\text{new}}(d_B, \lambda)$. By Theorem 1.3 and (6.2) we have $\Theta(\widetilde{f}_j) \not\equiv 0$ if and only if $\widetilde{f}_j(z_0) \neq 0$. By choosing $z_0 \in \mathcal{H}$ such that this holds, we get by (6.8), (6.9) and (6.1):

$$\widetilde{\epsilon}_j = \epsilon_j, \quad \text{and} \quad \widetilde{\tau}_j(n) = \tau_j(n) \quad \text{for all } n \in \mathbb{Z}^+. \tag{6.10}$$

Also, by (6.9), $\Theta : A(\Gamma_R \setminus \mathcal{H}, \lambda) \to A_{\text{new}}(d_B, \lambda)$ is a bijection if and only if $\Theta(\widetilde{f}_j) \not\equiv 0$ for all $j \in \{1, ..., D\}$, i.e. if and only if $\widetilde{f}_j(z_0) \neq 0$ for all $j \in \{1, ..., D\}$. Now Theorem 1.8 is proved.

[18] We can't apply Lemma 4.9, since we don't know $\epsilon_j = \widetilde{\epsilon}_j$ at this point.

□

Of course, by repeating the above procedure for each eigenvalue $\lambda$, we also obtain a proof of Theorem 1.6. In details: We choose a Hecke eigenbasis in each eigenspace $A(\Gamma_R \setminus \mathcal{H}, \lambda)$. We then obtain (changing our notation slightly from above) an orthonormal set of functions $\{\widetilde{f}_1, \widetilde{f}_2, \widetilde{f}_3, ...\} \subset L_2(\Gamma_R \setminus \mathcal{H})$ such that:

$$\begin{aligned} &\widetilde{f}_j \in A(\Gamma_R \setminus \mathcal{H}, \lambda_j), \qquad 0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq ... \\ &\mathcal{P}\widetilde{f}_j = \widetilde{\epsilon}_j \cdot \widetilde{f}_j, \qquad \widetilde{T}_n \widetilde{f}_j = \widetilde{\tau}_j(n) \cdot \widetilde{f}_j, \quad \forall n \in \mathbb{Z}^+ \\ &\{\widetilde{f}_j \mid \lambda_j = \lambda\} \text{ spans } A(\Gamma_R \setminus \mathcal{H}, \lambda), \text{ for each } \lambda > 0. \end{aligned} \tag{6.11}$$

Then by what we have proved, $\Theta : A(\Gamma_R \setminus \mathcal{H}, \lambda) \to A_{\text{new}}(d_B, \lambda)$ is a bijection for each $\lambda > 0$ if and only if:

$$\widetilde{f}_j(z_0) \neq 0 \qquad \forall j \in \{1, 2, ...\}. \tag{6.12}$$

The set of $z_0 \in \mathcal{H}$ *not* satisfying this is clearly of (Lebesgue) measure zero, since it is a countable union of regular arcs (use [H4, pp. 16–17, Remark 4.4, Prop. 4.5] together with compactness and an obvious extension of [O2, Ch. 13, §5 , Satz 4]).

To see that there is no mystery with the condition (6.12), notice that the choice of the functions $\widetilde{f}_1, \widetilde{f}_2, \widetilde{f}_3, ...$ in (6.11) is *unique*, except for permutations within each set $\{j \mid \lambda_j = \lambda\}$ and multiplication by scalars (of modulus one). Indeed, fixing $\lambda > 0$ with $A(\Gamma_R \setminus \mathcal{H}, \lambda) \neq \{0\}$, and fixing $z_0 \in \mathcal{H}$ so that $\Theta : A(\Gamma_R \setminus \mathcal{H}, \lambda) \to A_{\text{new}}(d_B, \lambda)$ is a bijection, then, by what we have proved, *every* Hecke basis of $A(\Gamma_R \setminus \mathcal{H}, \lambda)$ has to consist exactly of the inverse images under $\Theta$ of the newforms in $A_{\text{new}}(d_B, \lambda)$, multiplied by certain individual nonzero constants.

□

We finally prove the "multiplicity one theorem", Theorem 1.9. We use the notation in (6.11), and we also define $\widetilde{f}_0$ to be the normalized constant function on $\Gamma_R \setminus \mathcal{H}$; $\widetilde{f}_0 := \mu(\Gamma_R \setminus \mathcal{H})^{-1/2}$. Then $\{\widetilde{f}_0, \widetilde{f}_1, \widetilde{f}_2, ...\}$ forms an orthonormal and *complete* set of functions in $L_2(\Gamma_R \setminus \mathcal{H})$ (cf. [H3, p. 3], [H4, p. 316(lines 2-8)]).

We also define $\widetilde{\tau}_0(n)$ as in Remark 5.4, so that $\widetilde{T}_n \widetilde{f}_0 = \widetilde{\tau}_0(n) \cdot \widetilde{f}_0$.

We now claim:

(6.13) If $j_1, j_2 \in \{0, 1, 2, ...\}$, $j_1 \neq j_2$, then there are infinitely many primes $p$ such that $\widetilde{\tau}_{j_1}(p) \neq \widetilde{\tau}_{j_2}(p)$.

In fact, if $j_1, j_2 \geq 1$, then (6.13) follows from Theorem 1.8 and Theorem 4.6(d). In the case $j_1 = 0$, $j_2 \geq 1$, (6.13) follows from the fact that $\widetilde{\tau}_{j_2}(p) = O(p^{2/5+\epsilon})$ as $p \to \infty$ (through all primes), by Theorem 1.8 and Lemma 4.10, while $\widetilde{\tau}_{j_1}(p) = (p+1)/\sqrt{p}$ for all primes $p$ not dividing $d_B$, according to Remark 5.4.

From (6.13), Theorem 1.9 follows at once: Simply expand the two given functions $g_1, g_2$ in the $L_2$-basis $\{\widetilde{f}_0, \widetilde{f}_1, \widetilde{f}_2, ...\}$; then by (6.13), only *one* basis element can occur with nonzero coefficient, and this has to be the *same* basis element for $g_1$ as for $g_2$.

## 7. Some remarks on Siegel's theta function

In this section we give a *direct*, "trace-free" proof of part of Theorem 1.6, using properties of Siegel's theta function. Cf. [H5, p. 187(line 9)]. More precisely, we will prove the fact that $\Theta(\varphi)$ *belongs to the newspace* $A_{new}(d_B, \lambda)$, for any nonconstant eigenfunction $\varphi \in A(\Gamma_R \setminus \mathcal{H}, \lambda)$. Of course, by Lemma 6.1 and Theorem 1.3, an immediate consequence of this is that if $\varphi$ is an element of a Hecke basis in $A(\Gamma_R \setminus \mathcal{H})$, then $\Theta(\varphi)$ is a constant multiple of a newform in $A_{\mathrm{new}}(d_B, \lambda)$.

We will not hesitate to restrict ourselves in this section to a special case (see (7.12) below) in which all the main ideas are adequately illustrated.

We start by recalling some definitions and formulas concerning Siegel's theta function from [H5, pp. 129–131]. The original papers are [S3, S4].

Take $S$ and $P$ to be real symmetric, nonsingular $N \times N-$matrices such that $P$ is positive definite and $PS^{-1}P = S$ (i.e. $P$ is a "majorant" for $S$). We also take $S$ to be "even", i.e. $s_{jj} \in 2\mathbb{Z}$ and $s_{ij} \in \mathbb{Z}$ for $1 \leq i, j \leq N$. We will keep $\tau \in \mathcal{H}$, and define

$$R := (\mathrm{Re}\,\tau)S + i(\mathrm{Im}\,\tau)P.$$

Also, for any $N \times N-$matrix $A$, we will use the customary notation

$$A[X] := X^t A X, \qquad \text{for } X \in \mathbb{R}^N.$$

Let:

$$F_\psi(\tau) = \sum_{m \in \mathbb{Z}^N} e^{\pi i R[m+\psi]}, \qquad \text{for } \psi \in S^{-1}\mathbb{Z}^N,\ \tau \in \mathcal{H}. \tag{7.1}$$

(since $P$ is positive definite, the sum is rapidly absolutely convergent). This function is commonly referred to as *Siegel's theta function* with characteristic $\psi$.

One easily checks that:

$$F_\psi(\tau+1) = e^{\pi i S[\psi]} F_\psi(\tau). \tag{7.2}$$

In [H5], formulas are given that tells how $F_\psi$ transforms under any modular substitution. Below, we will need a similar formula for slightly more general substitutions:

**Lemma 7.1.** *Take* $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{R})$ *such that* $c > 0$ *and* $a, c, d \in \mathbb{Z}$, *but* $b$ *need not be an integer. Notice that* $cb = ad - 1 \in \mathbb{Z}$. *Fix* $f \in \mathbb{Z}^+$ *as a divisor of* $c$ *such that* $fb \in \mathbb{Z}$, *or equivalently,* $c \mid f(ad-1)$. *Then:*

$$F_\psi\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{n/2}(c\overline{\tau}+d)^{\nu/2} \sum_{\substack{\omega \in S^{-1}\mathbb{Z}^N \\ \bmod f\mathbb{Z}^N}} \lambda(S\omega) G_\omega(\tau), \tag{7.3}$$

*where we take principal values,* $n$ *and* $\nu$ *are the positive and negative index of* $S$ ($n + \nu = N$ *since* $S$ *is nonsingular),*

$$G_\omega(\tau) = \sum_{m \in \mathbb{Z}^N} e^{\pi i R[fm+\omega]}, \tag{7.4}$$

*and*

$$\lambda(k) = e^{\frac{\pi i}{4}(\nu - n)} \left|\det(S)\right|^{-1/2} c^{-N/2} \times \sum_{\substack{g_0 \in \mathbb{Z}^N \\ \bmod c\mathbb{Z}^N}} \exp\left[\frac{\pi i}{c}\left(aS[g_0+\psi] + 2k^t(g_0+\psi) + dS^{-1}[k]\right)\right] \tag{7.5}$$

*for* $k \in \mathbb{Z}^N$. *The* $\lambda(k)$ *is invariant as follows:*

$$\lambda(k + fSl) = \lambda(k) \qquad \text{for every} \quad l \in \mathbb{Z}^N. \tag{7.6}$$

*Proof.* By applying the Poisson summation formula (use [H5, p. 130(lines 8-11)]), and imitating Siegel's procedure [S3, p. 25], we establish that

$$F_\psi\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{n/2}(c\overline{\tau}+d)^{\nu/2} \sum_{m \in \mathbb{Z}^N} \lambda(m) e^{\pi i R[S^{-1}m]}, \tag{7.7}$$

where $\lambda(k)$ is as in (7.5). We now prove (7.6). For $f = c$ this is trivial, since each term in the sum in (7.5) is invariant under $k \longrightarrow k + cSl$.

For a general $f$ as above, we have:

$$\begin{aligned}
&\frac{\pi i}{c}\Big(aS[g_0+\psi]+2(k+fSl)^t(g_0+\psi)+dS^{-1}[k+fSl]\Big)\\
&=\frac{\pi i}{c}\Big(aS[g_0+fdl+\psi]+2k^t(g_0+fdl+\psi)+dS^{-1}[k]\\
&\qquad+2f(1-ad)l^tS(g_0+\psi)+f^2d(1-ad)S[l]\Big).
\end{aligned}$$

But here the last two terms inside the paranthesis belong to $2c\mathbb{Z}$, and can thus be ignored when exponentiating. Using the substitution $g_0 = h_0 - fdl$, we arrive at (7.6).

Now in (7.7) we replace $m$ by $S(\omega+fm)$, where $\omega \in S^{-1}\mathbb{Z}^N \pmod{f\mathbb{Z}^N}$ and (again) $m \in \mathbb{Z}^N$. Using (7.6), we then get (7.3). $\square$

Notice that if $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$, $c>0$, then we can take $f=1$ in the above lemma, and thus get:

$$F_\psi\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{n/2}(c\overline{\tau}+d)^{\nu/2}\sum_{\substack{\omega\in S^{-1}\mathbb{Z}^N\\ \bmod \mathbb{Z}^N}}\lambda(S\omega)F_\omega(\tau). \tag{7.8}$$

This is [H5, (2.4)].

When $cS^{-1}$ is "even", (7.8) and (7.5) simplify considerably [H5, (2.7)]:

$$\begin{aligned}
F_\psi\left(\frac{a\tau+b}{c\tau+d}\right) =& e^{\frac{\pi i}{4}(\nu-n)}(c\tau+d)^{n/2}(c\overline{\tau}+d)^{\nu/2}\left|\det(S)\right|^{-\frac{1}{2}}\\
&\times c^{-N/2}e^{\pi i abS[\psi]}\sum_{\substack{g_0\in\mathbb{Z}^N\\ \bmod c\mathbb{Z}^N}} e^{\pi i\frac{a}{c}S[g_0]}F_{a\psi}(\tau).
\end{aligned} \tag{7.9}$$

In the computations below, we need to evaluate some Gauss sums. As usual, for integers $x, y$ (not both zero) we let $(x,y)$ denote their greatest common divisor. We will use the quadratic reciprocity symbol $\left(\frac{x}{y}\right)$ only in the case when $y$ is positive and odd, and then we define it as in [L1, p. 47, def. 19], i.e. Jacobi style. For $x$ odd, we also denote

$$\epsilon_x = \begin{cases} 1 & \text{if } x\equiv 1(4)\\ i & \text{if } x\equiv 3(4)\end{cases}.$$

For $a, b, c$ integers, $c \geq 1,\ a \neq 0$ we introduce the following Gauss sums:

$$G(a,b,c) := \sum_{k \bmod c} \exp\left[2\pi i \frac{1}{c}(ak^2 + bk)\right] \tag{7.10}$$

$$G(a,c) := G(a,0,c) = \sum_{k \bmod c} \exp\left[2\pi i \frac{a}{c} k^2\right] \tag{7.11}$$

Below, we will need an explicit formula for $G(a, b, c)$ in the case where $c$ is odd. This is:

**Lemma 7.2.** *Take $c \in \mathbb{Z}^+$ odd. Then if $(a, c) = 1$ :*

$$G(a,b,c) = \sqrt{c}\left(\frac{a}{c}\right)\epsilon_c \cdot \exp\left[-2\pi i \frac{1}{c}(b^2 \cdot 4^* a^*)\right].$$

*Here $x^*$ refers to the multiplicative inverse modulo $c$. If $(a, c) = d > 1$, then*

$$G(a,b,c) = \begin{cases} d \cdot G(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}) & \text{if } d \mid b \\ 0 & \text{if } d \nmid b \end{cases}.$$

*Proof.* We first take $b = 0$, $(a, c) = 1$. By referring to [N1, p. 187(exercise 122)] or [L2, pp. 86(QS 4), 90(middle)], we see that

$$G(a,c) = \sqrt{c} \cdot \left(\frac{a}{c}\right) \cdot \epsilon_c.$$

We now take $b \in \mathbb{Z}$ arbitrary, keeping $(a, c) = 1$. Using

$$ak^2 + bk \equiv a(k + b2^*a^*)^2 - b^2 4^* a^* \pmod{c}$$

one then gets

$$G(a,b,c) = \exp\left(-2\pi i \cdot \frac{1}{c} \cdot (b^2 4^* a^*)\right) \cdot G(a,c),$$

as asserted.

Finally, we take $(a, c) = d > 1$. In (7.10) we then let $k = j + \frac{c}{d} \cdot l$, $l = 0, 1, ..., d-1,\ j = 1, 2, ..., \frac{c}{d}$. Notice that

$$a\left(j + \frac{c}{d} \cdot l\right)^2 + b\left(j + \frac{c}{d} \cdot l\right) \equiv aj^2 + bj + b \cdot \frac{c}{d} \cdot l \pmod{c}, \qquad \forall j, l \in \mathbb{Z}.$$

Hence if $d \nmid b$, the sum over $l$ is 0, for each fixed $j$, and thus $G(a,b,c) = 0$. If $d \mid b$, the exponential only depends on $j$, not on $l$, and we get

$$\begin{aligned} G(a,b,c) &= d \cdot \sum_{j=1}^{c/d} \exp\left[2\pi i \frac{1}{c}(aj^2 + bj)\right] \\ &= d \cdot \sum_{j \bmod c/d} \exp\left[2\pi i \frac{1}{c/d}\left(\frac{a}{d}j^2 + \frac{b}{d}j\right)\right] = d \cdot G(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}), \end{aligned}$$

as asserted. □

We now make our choice of a quaternion algebra $B$ over $\mathbb{Q}$. We take $q$ and $r$ to be positive integers which are relatively prime and square free, and suppose that

$$\text{(7.12)} \qquad \begin{cases} (a) & qr > 1 \\ (b) & q \equiv 1 \pmod 4,\ r \text{ odd} \\ (c) & \left(\dfrac{r}{p}\right) = -1 \text{ for each prime } p \text{ dividing } q\ \cdot \\ (d) & \left(\dfrac{q}{p}\right) = -1 \text{ for each prime } p \text{ dividing } r \end{cases}$$

We then take $B$ to be the quaternion algebra over $\mathbb{Q}$ with basis $\{1, i, j, k\}$ such that $i^2 = q$, $j^2 = r$, $k = ij = -ji$. This is exactly the case studied in [H5, section VIII], with the extra assumption that $r$ be odd. From [H5, (3.1), pp. 152(i), 155] we see that $B$ is an indefinite division algebra, and that a maximal order in $B$ is given by

(7.13)

$$R = \mathbb{Z}\left[\xi_1, \xi_2, \xi_3, \xi_4\right], \quad \text{where } \xi_1 = 1,\ \xi_2 = \frac{1+i}{2},\ \xi_3 = j,\ \xi_4 = \frac{j+k}{2}.$$

The (reduced) discriminant of $B$ and of $R$ is $d_B = qr$ [H5, p. 156, bottom].

We also fix our isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$ as (cf. [H5, (3.1)]):

$$x_1 + x_2 i + x_3 j + x_4 k \mapsto \begin{pmatrix} x_1 + x_2\sqrt{q} & \sqrt{r}(x_3 + x_4\sqrt{q}) \\ \sqrt{r}(x_3 - x_4\sqrt{q}) & x_1 - x_2\sqrt{q} \end{pmatrix}.$$

With this choice, and with $\xi_1, \xi_2, \xi_3, \xi_4$ as in (7.13), we get for $\mathsf{B}$ and $S$ (see (1.3), (1.4)):
(7.14)
$$\mathsf{B} = \begin{pmatrix} 1 & \frac{1+\sqrt{q}}{2} & 0 & 0 \\ 0 & 0 & \sqrt{r} & \sqrt{r}\frac{1+\sqrt{q}}{2} \\ 0 & 0 & \sqrt{r} & \sqrt{r}\frac{1-\sqrt{q}}{2} \\ 1 & \frac{1-\sqrt{q}}{2} & 0 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & -2t & 0 & 0 \\ 0 & 0 & -2r & -r \\ 0 & 0 & -r & 2rt \end{pmatrix}$$

Here $t := (q-1)/4 \in \mathbb{Z}$. Cf. [H5, (3.1),(8.2)].

We will take $z, z_0 \in \mathcal{H}$ to be *fixed;* then $P = P_{zz_0}$ (see (1.5)) is a fixed majorant of $S$, and we can write the integration kernel function $\theta(z;\tau)$ simply as $\operatorname{Im}\tau \cdot F_0(\tau)$. Cf. (1.6), (7.1).

We make a few remarks on the matrix $S$ that will be useful later. First of all,

$$n = \nu = 2 \qquad \text{(positive and negative index of } S\text{)},$$

$$\det(S) = q^2r^2 \text{ and } S^{-1} = \frac{1}{qr}\begin{pmatrix} 2tr & r & 0 & 0 \\ r & -2r & 0 & 0 \\ 0 & 0 & -2t & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \tag{7.15}$$

and for any $k \in \mathbb{R}^4$ this easily gives

$$S^{-1}[k] = 2\Big(\frac{1}{4}k_1^2 - \frac{1}{4q}(2k_2-k_1)^2 - \frac{1}{4r}k_3^2 + \frac{1}{4qr}(2k_4-k_3)^2\Big). \tag{7.16}$$

The set $S\mathbb{Z}^4$ is given by

**Lemma 7.3.**

$$S\mathbb{Z}^4 = \{k = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} \in \mathbb{Z}^4 \mid \quad q \mid k_1 - 2k_2 \;\wedge\; r \mid k_3 \;\wedge\; qr \mid 2k_4 - k_3\}.$$

*Proof.* Clearly $S\mathbb{Z}^4 \subset \mathbb{Z}^4$. Take $k \in \mathbb{Z}^4$. $k \in S\mathbb{Z}^4$ holds if and only if $S^{-1}k \in \mathbb{Z}^4$. By (7.15), this is equivalent to

$$2trk_1 + rk_2 \equiv rk_1 - 2rk_2 \equiv -2tk_3 - k_4 \equiv -k_3 + 2k_4 \equiv 0 \pmod{qr},$$

i.e.

$$q \mid k_1 - 2k_2 \;\wedge\; q \mid 2tk_1 + k_2 \;\wedge\; qr \mid 2k_4 - k_3 \;\wedge\; qr \mid 2tk_3 + k_4.$$

But the condition $q \mid k_1 - 2k_2$ implies $2tk_1 + k_2 \equiv (4t+1)k_2 = qk_2 \equiv 0 \pmod{q}$. Similarly, if $qr \mid 2k_4 - k_3$, then $2(2tk_3 + k_4) \equiv qk_3 \pmod{qr}$, and thus $\big[qr \mid 2tk_3 + k_4 \Longleftrightarrow r \mid k_3\big]$. $\square$

The $\Gamma_0(d_R)$–invariance in Theorem 1.2 now takes the form

$$\text{Im }(T\tau)\cdot F_0(T\tau) = \text{Im }\tau \cdot F_0(\tau), \qquad \text{for each } T \in \Gamma_0(qr). \tag{7.17}$$

This is a special case of [BJ1, Prop. 4.1], but we give a proof here for our somewhat more explicit setting. (Cf. [H5, p. 158-159], where (7.17) is proved for $T \in \Gamma_0(4qr)$.)

**Lemma 7.4.** *The elements in* $M = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr) \mid c \text{ odd}\}$ *generates the full group* $\Gamma_0(qr)$.

*Proof.* Take $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr)$ with $c$ even. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a-qrb & b \\ c-qrd & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ qr & 1 \end{pmatrix},$$

and both factors in the right hand side lie in $M$. $\square$

Now take $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr)$. By the lemma, it is enough to consider the case where $c$ is odd. We may then also take $c > 0$.

From (7.15) we see that $cS^{-1}$ is "even", so (7.9) applies, giving

(7.18)

$$F_0\left(\frac{a\tau+b}{c\tau+d}\right) = |c\tau+d|^2(qr)^{-1}c^{-2}\cdot\left(\sum_{g_0\in\mathbb{Z}^4 \bmod c\mathbb{Z}^4} \exp(\pi i\frac{a}{c}S[g_0])\right) F_0(\tau).$$

But by (7.14):

$$\sum_{g_0\in\mathbb{Z}^4 \bmod c\mathbb{Z}^4} \exp(\pi i\frac{a}{c}S[g_0]) = \left(\sum_{n_1,n_2 \pmod c} \exp(2\pi i\frac{a}{c}(n_1^2+n_1n_2-tn_2^2))\right)$$
$$\times\left(\sum_{n_3,n_4 \pmod c} \exp(2\pi i\frac{a}{c}(-rn_3^2-rn_3n_4+rtn_4^2))\right).$$

Here we substitute (remember $2 \nmid c$)

$$\langle n_1,n_2,n_3,n_4\rangle \equiv \langle m_1-m_2, 2m_2, m_3-m_4, 2m_4\rangle \pmod c$$

and get

$$= G(a,c) \cdot G(-aq,c) \cdot G(-ar,c) \cdot G(arq,c).$$

We now use Lemma 7.2 and quadratic reciprocity (remember $(a,c) = 1$, $rq \mid c$ and $q \equiv 1(4)$):

$$\begin{aligned}
&= G(a,c) \cdot q \cdot G(-a, \frac{c}{q}) \cdot r \cdot G(-a, \frac{c}{r}) \cdot rq \cdot G(a, \frac{c}{rq}) \\
&= c^2 rq \left(\frac{a}{c}\right) \left(\frac{-a}{c/q}\right) \left(\frac{-a}{c/r}\right) \left(\frac{a}{c/rq}\right) \epsilon_c \epsilon_{c/q} \epsilon_{c/r} \epsilon_{c/rq} \\
&= c^2 qr \left(\frac{a}{qr}\right) \left(\frac{-a}{rq}\right) \cdot \left\{\begin{matrix} 1 & \text{if } r \equiv 1(4) \\ -1 & \text{if } r \equiv 3(4) \end{matrix}\right\} \\
&= c^2 rq.
\end{aligned}$$

Inserting this in (7.18), we get (7.17).

□

The following lemma gives a bound on $\theta(z;\tau) = \operatorname{Im}\tau \cdot F_0(\tau)$ as $\tau$ approaches any cusp of $\Gamma_0(qr)$ cuspidally (for each such cusp is equivalent to $i\infty$ under $\mathrm{SL}(2,\mathbb{Z})$).

**Lemma 7.5.** *For each fixed $M \in SL(2,\mathbb{Z})$ :*

$$\theta(z; M\tau) = O(\operatorname{Im}\tau), \qquad \text{as } \operatorname{Im}\tau \to \infty, \text{ uniformly in } z \in \mathcal{H}. \tag{7.19}$$

*Proof.* Write $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We may assume $c \geq 0$.

By the $\Gamma_R$-invariance of $\theta(z;\tau)$ (Theorem 1.2) it is enough to consider $z$ restricted to a fundamental polygon $\mathcal{F}_{\Gamma_R}$ for $\Gamma_R \setminus \mathcal{H}$. $\mathcal{F}_{\Gamma_R}$ is compact, and hence we can bound $P_{zz_0}$ (see (1.5)) from below: There is some $a > 0$ such that $k^t P_{zz_0} k \geq a \cdot k^t k$ for all $k \in \mathbb{R}^4$, $z \in \mathcal{F}_{\Gamma_R}$. In the case $c > 0$, we use this and (7.1) to check that in (7.8), each $F_\omega(\tau)$ in the right hand side tends either to 0 or 1 as $\operatorname{Im}\tau \to \infty$, uniformly over $z \in \mathcal{F}_{\Gamma_R}$. From this, (7.19) follows. If $c = 0$, we use the same bound on $P_{zz_0}$ to prove (7.19) directly from (7.1). □

The fact that the correspondence $\Theta$ maps *into the newspace* on $\Gamma_0(qr)$ will be deduced as a consequence of the following two relations, (7.21) and (7.24), satisfied by the function $F_0(\tau)$ as we have chosen it.

**Theorem 7.6.** *If $p$ is any prime dividing $qr$, and $\alpha_j \in \Gamma_0\left(\frac{qr}{p}\right)$ are coset representatives such that*

$$\Gamma_0(\frac{qr}{p}) = \bigsqcup_j \Gamma_0(qr)\alpha_j, \tag{7.20}$$

*then*

$$\sum_j (\operatorname{Im} \alpha_j\tau) F_0(\alpha_j\tau) = 0, \qquad \forall\tau \in \mathcal{H}. \tag{7.21}$$

The second relation involves the group

$$\begin{aligned} G_p &= \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}^{-1} \Gamma_0\left(\frac{qr}{p}\right) \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \\ &= \{\begin{pmatrix} a & b/p \\ cp & d \end{pmatrix} \mid \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0\left(\frac{qr}{p}\right)\}, \end{aligned} \tag{7.22}$$

where again $p$ is a prime divisor of $qr$. Notice that $\Gamma_0(qr) \subset G_p$.

**Theorem 7.7.** *If $p$ is any prime dividing $qr$, and $\beta_j \in G_p$ are coset representatives such that*

$$G_p = \bigsqcup_j \Gamma_0(qr)\beta_j, \tag{7.23}$$

*then*

$$\sum_j (\operatorname{Im} \beta_j\tau) F_0(\beta_j\tau) = 0, \qquad \forall\tau \in \mathcal{H}. \tag{7.24}$$

Before proving these two relations, we show how they lead to the desired conclusion about our spectral correspondence $\Theta$. Take $\lambda > 0$ and let $\varphi \in A(\Gamma_R \setminus \mathcal{H}, \lambda)$ be given. We will then prove that $\Theta(\varphi) \in A_{\text{new}}(qr, \lambda)$.

Recall (1.17) and Definitions 1.4, 1.5.

**Lemma 7.8.** *For any $N \in \mathbb{Z}^+$, $\lambda > 0$, $A_{old}(N, \lambda)$ is spanned by the spaces in*

$$\{A(\frac{N}{p}, \lambda) \mid p \text{ prime, } p \mid N\} \quad \bigcup \quad \{\delta_p A(\frac{N}{p}, \lambda) \mid p \text{ prime, } p \mid N\}.$$

*Proof.* Take $M, l \in \mathbb{Z}^+$ as in Definition 1.4, i.e. assume $lM \mid N$ and $M \neq N$. Then if $lM < N$ there is a prime $p$ dividing $N/lM$, and, by Lemma 4.4(e),

$$\delta_l A(M, \lambda) \subset A(Ml, \lambda) \subset A(N/p, \lambda),$$

since $lM \mid N/p$. In the remaining case, $lM = N$, we must have $l > 1$, and then for any prime $p$ dividing $l$ :

$$\delta_l A(M,\lambda) = \delta_p \delta_{l/p} A(M,\lambda) \subset \delta_p A(Ml/p,\lambda) = \delta_p A(N/p,\lambda).$$

□

By Theorem 1.3, $\Theta(\varphi) \in A(qr,\lambda)$. In view of Lemma 7.8 and Definition 1.5, to prove $\Theta(\varphi) \in A_{\mathrm{new}}(qr,\lambda)$, it is enough to prove that for each prime $p$ dividing $qr$, and each $g \in A(qr/p,\lambda)$, $\Theta(\varphi)$ is orthogonal both to $g$ and to $\delta_p g$ in $A(qr,\lambda)$. Let $\mathcal{F}$ be a (nice) fundamental domain for the group $\Gamma_0(qr/p)$. Then by (7.20), $\bigsqcup_j \alpha_j \mathcal{F}$ is a fundamental domain for $\Gamma_0(qr)$, and the inner product $\langle \Theta(\varphi), g\rangle$ can be written as

$$\begin{aligned}
&\sum_j \int_{\alpha_j \mathcal{F}} \Theta(\varphi)(\tau) \cdot \overline{g(\tau)}\, d\mu(\tau) \\
&= \sum_j \int_{\alpha_j \mathcal{F}} \left( \int_{\mathcal{F}_{\Gamma_R}} \theta(z;\tau)\varphi(z)\, d\mu(z) \right) \cdot \overline{g(\tau)}\, d\mu(\tau) \\
&= \int_{\mathcal{F}_{\Gamma_R}} \varphi(z) \cdot \left( \sum_j \int_{\alpha_j \mathcal{F}} \theta(z;\tau)\overline{g(\tau)}\, d\mu(\tau) \right) d\mu(z) \\
&\qquad \left\{ \begin{array}{l} \text{Using Lemma 7.5, } \Gamma_0(qr)\text{-invariance, and the fact that} \\ g(\tau) \text{ is a cusp form, we see that } \theta(z;\tau)\varphi(z)\overline{g(\tau)} \text{ is} \\ \text{bounded for } (\tau,z) \in \mathcal{H}\times\mathcal{H}. \text{ This justifies the change} \\ \text{of order of integration.} \end{array} \right\} \\
&= \int_{\mathcal{F}_{\Gamma_R}} \varphi(z) \cdot \int_{\mathcal{F}} \left( \sum_j \theta(z;\alpha_j \tau) \right) \overline{g(\tau)}\, d\mu(\tau)\, d\mu(z) \\
&\qquad \left\{ \text{since } g \text{ is } \Gamma_0(qr/p)\text{-invariant} \right\} \\
&= 0. \qquad \{\text{by (7.21)}\}
\end{aligned}$$

The proof that $\Theta(\varphi)$ is also orthogonal to $\delta_p g$ is very similar: This time we use (7.23) to write our integral as $\sum_j \int_{\beta_j \mathcal{F}'}$, where $\mathcal{F}'$ is a fundamental domain for the group $G_p$, and then use that $\delta_p g$ is invariant under $G_p$ (which is straightforward from our definitions (1.17) and (7.22)), and the relation (7.24).

This completes the proof of $\Theta(\varphi) \in A_{\mathrm{new}}(qr,\lambda)$.

□

We now prove Theorem 7.6. Fix a prime divisor $p$ of $qr$. By (7.17) it is enough to prove (7.21) for *one* choice of the representatives $\alpha_j$.

**Lemma 7.9.** *There are* $\alpha_j = \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} \in \Gamma_0\left(\frac{qr}{p}\right)$ *for* $j = 0, 1, ..., (p-1)$ *such that*

$$\begin{aligned} \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} &\equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{\frac{qr}{p}}, \\ \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} &\equiv \begin{pmatrix} 0 & -1 \\ 1 & j \end{pmatrix} \pmod{p} \\ a_j \neq 0,\ c_j > 0,&\ 2 \nmid c_j,\ 4 \mid d_j. \end{aligned} \tag{7.25}$$

*We also let* $\alpha_p = \begin{pmatrix} a_p & b_p \\ c_p & d_p \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. *For any such choice of matrices* $\alpha_0, ..., \alpha_p$ *we have*

$$\Gamma_0\left(\frac{qr}{p}\right) = \bigsqcup_{j=0}^{p} \Gamma_0(qr)\alpha_j. \tag{7.26}$$

*Proof.* Take $j \in \{0, 1, ..., p-1\}$. Let $N = 4qr$. Notice that $p$ and $qr/p$ are odd and relatively prime, by our assumptions on $q$ and $r$. By the Chineese remainder theorem, there is a unique matrix $M \in \mathrm{SL}(2, \mathbb{Z}/N\mathbb{Z})$ such that $M \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{qr/p}$, $M \equiv \begin{pmatrix} 0 & -1 \\ 1 & j \end{pmatrix} \pmod{p}$, and – say – $M \equiv \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \pmod{4}$. By [M2, p. 104, Th. 4.2.1(1)], there is a matrix $\alpha_j = \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$ that is congruent to $M$ modulo $N$. By examining the proof in [M2, p. 104], it is easily seen that we can also take $c_j > 0$. Then $\alpha_j$ fulfills all conditions above.

The proof of (7.26) only uses $\alpha_j \in \Gamma_0(qr/p)$ and the congruence class of $\alpha_j$ modulo $p$. Take $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr/p)$. We will prove that there is *exactly one* $j \in \{0, 1, ..., p\}$ such that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr)\alpha_j$. Notice:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(qr)\alpha_j &\iff \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix}^{-1} \in \Gamma_0(qr) \\ &\iff \begin{pmatrix} * & * \\ cd_j - dc_j & * \end{pmatrix} \in \Gamma_0(qr). \end{aligned}$$

This is equivalent to $cd_j - dc_j \equiv 0 \pmod p$, since the matrix a priori belongs to $\Gamma_0(qr/p)$. But since $\langle c_j, d_j\rangle \equiv \langle 1, j\rangle \pmod p$ for $j \in \{0, .., p-1\}$ and $\langle c_p, d_p\rangle = \langle 0, 1\rangle$, $cd_j - dc_j \equiv 0 \pmod p$ clearly holds for exactly one $j \in \{0, 1, ..., p\}$, as asserted. □

We choose our $\alpha_j$'s as in the lemma. Applying (7.8) for $\alpha_0, ...\alpha_{p-1}$, we see that (7.21) would follow if we could show

$$\sum_{j=0}^{p-1} \lambda_{\alpha_j}(k) = \begin{cases} 0 & \text{if } k \notin S\mathbb{Z}^4 \\ -1 & \text{if } k \in S\mathbb{Z}^4 \end{cases}, \tag{7.27}$$

where for $j = 0, 1, ..., p-1$ (cf. (7.5), (7.14) and (7.15))

(7.28)

$$\lambda_{\alpha_j}(k) = \frac{1}{c_j^2 qr} \cdot \exp(\pi i \frac{d_j}{c_j} S^{-1}[k])$$
$$\times \sum_{n_1, n_2 \pmod{c_j}} \exp\left(2\pi i \frac{1}{c_j}(a_j(n_1^2 + n_1 n_2 - t n_2^2) + k_1 n_1 + k_2 n_2)\right)$$
$$\times \sum_{n_3, n_4 \pmod{c_j}} \exp\left(2\pi i \frac{1}{c_j}(-a_j r(n_3^2 + n_3 n_4 - t n_4^2) + k_3 n_3 + k_4 n_4)\right).$$

As before, we substitute

$$\langle n_1, n_2, n_3, n_4\rangle \equiv \langle m_1 - m_2, 2m_2, m_3 - m_4, 2m_4\rangle \pmod{c_j}$$

in (7.28) (remember $2 \nmid c_j$), and get

$$\lambda_{\alpha_j}(k) = \frac{1}{c_j^2 qr} \exp(\pi i \frac{d_j}{c_j} S^{-1}[k]) \cdot G(a_j, k_1, c_j) \cdot G(-a_j q, 2k_2 - k_1, c_j)$$
$$\times G(-a_j r, k_3, c_j) \cdot G(a_j qr, 2k_4 - k_3, c_j).$$

Now the computation splits into the cases $p \mid q$ and $p \mid r$. We will here assume $p \mid q$. The case $p \mid r$ only requires very minor changes to what follows, and is left to the reader.

Now by (7.25), $(a_j, c_j) = 1$, $(-a_j q, c_j) = q/p$, $(-a_j r, c_j) = r$ and $(a_j qr, c_j) = qr/p$. Hence by Lemma 7.2

$$\lambda_{\alpha_j}(k) = 0 \qquad \text{if } \frac{q}{p} \nmid (2k_2 - k_1) \text{ or } r \nmid k_3 \text{ or } \frac{qr}{p} \nmid (2k_4 - k_3).$$

Hence (7.27) is proved except when $\frac{q}{p} \mid (2k_2 - k_1)$, $r \mid k_3$ and $\frac{qr}{p} \mid (2k_4 - k_3)$ (cf. Lemma 7.3). In that case, we use (7.16) and Lemma 7.2, to see that

$$\lambda_{\alpha_j}(k) = \frac{qr}{c_j^2p^2}\exp(\pi i\frac{d_j}{c_j}S^{-1}[k])\cdot G(a_j,k_1,c_j)\cdot G(-a_jp,\frac{2k_2-k_1}{q/p},\frac{c_j}{q/p})$$

$$\times G(-a_j,\frac{k_3}{r},\frac{c_j}{r})\cdot G(a_jp,\frac{2k_4-k_3}{qr/p},\frac{c_j}{qr/p})$$

$$= \frac{1}{p}\exp\left[2\pi i\frac{1}{c_j}\left(k_1^2\frac{1}{4}d_j\right)\right]\exp\left[-2\pi i\frac{1}{c_jp^2/q}\left(\left(\frac{2k_2-k_1}{q/p}\right)^2\frac{1}{4}d_j\right)\right]$$

$$\times\exp\left[-2\pi i\frac{1}{c_j/r}\left(\left(\frac{k_3}{r}\right)^2\frac{1}{4}d_j\right)\right]\exp\left[2\pi i\frac{1}{c_jp^2/qr}\left(\left(\frac{2k_4-k_3}{qr/p}\right)^2\frac{1}{4}d_j\right)\right]$$

$$\times\left(\frac{a_j}{c_j}\right)\left(\frac{-a_jp}{c_jp/q}\right)\left(\frac{-a_j}{c_j/r}\right)\left(\frac{a_jp}{c_jp/qr}\right)\epsilon_{c_j}\epsilon_{c_jp/q}\epsilon_{c_j/r}\epsilon_{c_jp/qr}$$

$$\times\exp\left[-2\pi i\frac{1}{c_j}(k_1^24^*a_j^*)\right]\exp\left[2\pi i\frac{1}{c_jp/q}\left(\left(\frac{2k_2-k_1}{q/p}\right)^24^*(a_jp)^*\right)\right]$$

$$\times\exp\left[2\pi i\frac{1}{c_j/r}\left(\left(\frac{k_3}{r}\right)^24^*a_j^*\right)\right]\exp\left[-2\pi i\frac{1}{c_jp/qr}\left(\left(\frac{2k_4-k_3}{qr/p}\right)^24^*(a_jp)^*\right)\right],$$

where $x^*$ refers to inverse modulo $c_j$, $c_jp/q$, $c_j/r$ and $c_jp/qr$, respectively. But $a_j^* \equiv d_j \pmod{c_j}$, and by (7.25) we have $a_j \equiv 0$, $d_j \equiv j$ and $c_j \equiv 1 \pmod p$, and $4 \mid d_j$. Using this, the product of the eight exponentials can be shown to simplify into

$$\exp\left[2\pi i\frac{1}{p}\left(-(c_jp/q)^*\left(\frac{2k_2-k_1}{q/p}\right)^24^*d_j+(c_jp/qr)^*\left(\frac{2k_4-k_3}{qr/p}\right)^24^*d_j\right)\right]$$

$$= \exp\left[2\pi i\frac{j}{p}(4qr/p)^*[-r(2k_2-k_1)^2+(2k_4-k_3)^2]\right],$$

where $x^*$ refers to inverse modulo $p$. The product of reciprocity symbols and $\epsilon$'s simplifies as

$$\left(\frac{-1}{qr/p}\right)\left(\frac{p}{r}\right)\begin{cases}1 & \text{if } \frac{q}{p}\equiv r\equiv 1(4)\\ -1 & \text{otherwise}\end{cases},$$

which by reciprocity and (7.12 b,c) is $-1$. Hence

$$\lambda_{\alpha_j}(k) = -\frac{1}{p}\exp\left[2\pi i\frac{j}{p}\left((4qr/p)^*[-r(2k_2-k_1)^2+(2k_4-k_3)^2]\right)\right].$$

Thus:

$$\sum_{j=0}^{p-1} \lambda_{\alpha_j}(k) = \begin{cases} -1 & \text{if } -r(2k_2-k_1)^2 + (2k_4-k_3)^2 \equiv 0 \pmod p \\ 0 & \text{otherwise} \end{cases}.$$

But notice that, since $\left(\frac{r}{p}\right) = -1$ :

$$\begin{aligned}
& -r(2k_2-k_1)^2 + (2k_4-k_3)^2 \equiv 0 \pmod p \\
& \iff 2k_2 - k_1 \equiv 2k_4 - k_3 \equiv 0 \pmod p \\
& \iff q \mid 2k_2 - k_1, \quad r \mid k_3, \quad qr \mid 2k_4 - k_3 \\
& \qquad \Big\{\text{Since we are already assuming } \frac{q}{p} \mid 2k_2 - k_1,\ r \mid k_3,\ \frac{qr}{p} \mid 2k_4 - k_3.\Big\} \\
& \iff k \in S\mathbb{Z}^4. \qquad \Big\{\text{By Lemma 7.3.}\Big\}
\end{aligned}$$

Thus (7.27) is proved, and hence also Theorem 7.6.

□

We now prove Theorem 7.7. Fix a prime divisor $p$ of $qr$. Again by (7.17) it is enough to prove (7.24) for *one* choice of the representatives $\beta_j$.

**Lemma 7.10.** *There are* $\begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} \in \Gamma_0\left(\frac{qr}{p}\right)$ *for* $j = 0, 1, ..., (p-1)$ *such that*

$$\begin{aligned}
&\begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{\frac{qr}{p}}, \\
&\begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix} \equiv \begin{pmatrix} j & 1 \\ -1 & 0 \end{pmatrix} \pmod p \\
&a_j \neq 0,\ c_j > 0,\ 2 \nmid c_j,\ 4 \mid d_j.
\end{aligned} \tag{7.29}$$

*We then let* $\beta_j = \begin{pmatrix} a_j & b_j/p \\ c_j p & d_j \end{pmatrix} \in G_p$ *for* $j = 0, 1, ..., p-1$ *(cf. (7.22)) and* $\beta_p = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in G_p$*. For any such choice of matrices* $\beta_0, ..., \beta_p$ *we have*

$$G_p = \bigsqcup_{j=0}^{p} \Gamma_0(qr)\beta_j. \tag{7.30}$$

*Proof.* Very similar to the proof of Lemma 7.9. In the proof of (7.30), notice that for $\begin{pmatrix} a & b/p \\ cp & d \end{pmatrix} \in G_p$ and $j \in \{0, 1, ..., p\}$, the product $\begin{pmatrix} a & b/p \\ cp & d \end{pmatrix} \beta_j^{-1} = \begin{pmatrix} * & (ba_j - ab_j)/p \\ * & * \end{pmatrix}$ belongs to $G_p$, and hence it lies in $\Gamma_0(qr)$ if and only if $ba_j - ab_j \equiv 0 \pmod{p}$. □

Take $\beta_j$ as in the lemma, and keep $j \in \{0, 1, ..., p-1\}$. We apply Lemma 7.1 with $f = p$ :

$$(\mathrm{Im}\, \beta_j \tau)\, F_0(\beta_j \tau) = (\mathrm{Im}\, \tau) \cdot \sum_{\substack{\omega \in S^{-1}\mathbb{Z}^4 \\ \bmod p\mathbb{Z}^4}} \lambda_{\beta_j}(S\omega) G_\omega(\tau), \tag{7.31}$$

where $G_\omega(\tau)$ and $\lambda_{\beta_j}(k)$ (for $k \in \mathbb{Z}^4$) are given by (7.4) and (7.5) respectively, applied to the matrix $\beta_j$, and with $f = p$.

We now compute $\lambda_{\beta_j}(k)$.

$$\begin{aligned}
\lambda_{\beta_j}(k) &= \frac{1}{c_j^2 p^2 qr} \exp\left[\pi i \frac{d_j}{c_j p} S^{-1}[k]\right] \sum_{g \in \mathbb{Z}^4 \pmod{c_j p \mathbb{Z}^4}} \exp\left[\pi i \frac{a_j}{c_j p} S[g] + 2\pi i \frac{1}{c_j p} k^t g\right] \\
&= \frac{1}{c_j^2 p^2 qr} \exp\left[\pi i \frac{d_j}{c_j p} S^{-1}[k]\right] \\
&\times \sum_{n_1, n_2 \pmod{c_j p}} \exp\left[2\pi i \frac{1}{c_j p}(a_j(n_1^2 + n_1 n_2 - t n_2^2) + k_1 n_1 + k_2 n_2)\right] \\
&\times \sum_{n_3, n_4 \pmod{c_j p}} \exp\left[2\pi i \frac{1}{c_j p}(-a_j r(n_3^2 + n_3 n_4 - t n_4^2) + k_3 n_3 + k_4 n_4)\right].
\end{aligned}$$

We substitute $\langle n_1, n_2, n_3, n_4 \rangle \equiv \langle m_1 - m_2, 2m_2, m_3 - m_4, 2m_4 \rangle \pmod{c_j p}$ (remember $2 \nmid c_j p$), and get

$$\begin{aligned}
\lambda_{\beta_j}(k) &= \frac{1}{c_j^2 p^2 qr} \exp\left[\pi i \frac{d_j}{c_j p} S^{-1}[k]\right] \cdot G(a_j, k_1, c_j p) \cdot G(-a_j q, 2k_2 - k_1, c_j p) \\
&\times G(-a_j r, k_3, c_j p) \cdot G(a_j qr, 2k_4 - k_3, c_j p).
\end{aligned}$$

The computation of $\lambda_{\beta_j}(k)$ from here on is very similar to our computation in the proof of Theorem 7.6, and we simply state the result.

We first assume $p \mid q$. Then:

For $j \in \{1, ..., p-1\}$ :

$$\lambda_{\beta_j}(k) = \begin{cases} \exp\left[2\pi i \frac{1}{p}(4rj)^*(rk_1^2 - k_3^2)\right] & \text{if } k \in S\mathbb{Z}^4 \\ 0 & \text{otherwise} \end{cases}$$

For $j = 0$ :

$$\lambda_{\beta_0}(k) = \begin{cases} -p & \text{if } k \in S\mathbb{Z}^4 \ \wedge \ k_1 \equiv k_3 \equiv 0 \pmod p \\ 0 & \text{otherwise} \end{cases}$$

Here $x^*$ refers to inverse modulo $p$.

By (7.12 c), we have $rk_1^2 - k_3^2 \equiv 0 \pmod p$ if and only if $k_1 \equiv k_3 \equiv 0 \pmod p$. Thus

$$\sum_{j=1}^{p-1} \lambda_{\beta_j}(k) = \begin{cases} p-1 & \text{if } k \in S\mathbb{Z}^4 \ \wedge \ k_1 \equiv k_3 \equiv 0 \pmod p \\ -1 & \text{if } k \in S\mathbb{Z}^4 \ \wedge \ \neg\, (k_1 \equiv k_3 \equiv 0 \pmod p) \\ 0 & \text{if } k \notin S\mathbb{Z}^4 \end{cases},$$

and hence

$$\sum_{j=0}^{p-1} \lambda_{\beta_j}(k) = \begin{cases} -1 & \text{if } k \in S\mathbb{Z}^4 \\ 0 & \text{otherwise} \end{cases}. \tag{7.32}$$

The same relation holds in the case $p \mid r$, and the proof is very similar. Assuming $p \mid r$, the computation of $\lambda_{\beta_j}(k)$ gives:

For $j \in \{1, ..., p-1\}$ :

$$\lambda_{\beta_j}(k) = \begin{cases} \exp\left[2\pi i \frac{1}{p}(4qj)^*(qk_1^2 - (2k_2 - k_1)^2)\right] & \text{if } k \in S\mathbb{Z}^4 \\ 0 & \text{otherwise} \end{cases}$$

For $j = 0$ :

$$\lambda_{\beta_0}(k) = \begin{cases} -p & \text{if } k \in S\mathbb{Z}^4 \ \wedge \ k_1 \equiv k_2 \equiv 0 \pmod p \\ 0 & \text{otherwise} \end{cases}$$

Here $x^*$ refers to inverse modulo $p$. By (7.12 d), $qk_1^2 - (2k_2 - k_1)^2 \equiv 0 \pmod p$ if and only if $k_1 \equiv (2k_2 - k_1) \equiv 0 \pmod p$, i.e. if and only if $k_1 \equiv k_2 \equiv 0 \pmod p$. Hence (7.32) follows as before.

From (7.31) and (7.32) we conclude:

$$\sum_{j=0}^{p-1} (\operatorname{Im} \beta_j \tau)\, F_0(\beta_j \tau) = -\,(\operatorname{Im} \tau) \cdot \sum_{\substack{\omega \in \mathbb{Z}^4 \\ \bmod p\mathbb{Z}^4}} G_\omega(\tau) = -\,(\operatorname{Im} \tau) \cdot F_0(\tau),$$

using the definitions (7.1) and (7.4). This proves (7.24), and hence Theorem 7.7. □

## References


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

[BV] E. Balslev and A. Venkov, The Weyl law for subgroups of the modular group, Geom. Funct. Anal. **8** (1998), 437–465.

[BR] J. Bernstein and A. Reznikov, Analytic continuation of representations and estimates of automorphic forms, Ann. of Math. **150** (1999), 329–352.

[BJ1] J. Bolte, S. Johansson, Theta-lifts of Maass waveforms, in *Emerging Applications of Number Theory,* (edited by D. A. Hejhal, J. Friedman, M. Gutzwiller, A. Odlyzko), IMA Vol. Math. Appl. **109**, Springer-Verlag, New York, 1999, 39–72.

[BJ2] J. Bolte, S. Johansson, A Spectral Correspondence for Maaß waveforms, Geom. Funct. Anal., **9** (1999), 1128–1155.

[B] D. Bump, *Automorphic forms and representations,* Cambridge Univ. Press, 1997.

[BDHI] D. Bump, W. Duke, J. Hoffstein, and H. Iwaniec, An estimate for Hecke Eigenvalues of Maass Forms, Inter. Math. Res. Notices (1992), 75–81.

[BS] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press, New York, 1966.

[C1] W. Casselman, On some results of Atkin and Lehner, Math. Ann. **201** (1973), 301–314.

[C2] H. Cohen, *A Course in Computational Algebraic Number Theory,* Springer-Verlag, Berlin, 1993.

[C3] H. Cohn, *Advanced Number Theory,* Dover Publications Inc., 1980.

[CL] J. B. Conrey and Xian-Jin Li, On the trace of Hecke operators for Maass forms for congruence subgroups $\Gamma_0(N)$, preprint, [www.math.byu.edu/∼xianjin/papers.html].

[D] K. Dedekind, Über die Anzahl der Ideal-Klassen in den verschiedenen Ordnungen eines endlichen Körpers, in *Gesammelte matematische Werke,* Vol. I, Chelsea Publishing Co., New York, 1968, 105–158

[E] M. Eichler, *Lectures on Modular Correspondences,* Tata Institute, Bombay, 1956.

[G2] R. Gunning, *Lectures on Modular Forms,* Annals of Math. Studies No. 48, Princeton University Press, 1962.

[H1] E. Hecke, Über Modulfunktionen und die Dirichletscher Reihen mit Eulerscher Produktentwicklung I, II, Math. Ann. **114**, 1–28; 316–351.

[H2] D. A. Hejhal, The Selberg trace formula and the Riemann zeta function, Duke Math. J. **43** (1976), 441–482.

[H3] D. A. Hejhal, *The Selberg Trace Formula for* $PSL(2,\mathbb{R})$, Vol. 1, Lecture Notes in Math. **548**, Springer-Verlag, Berlin, 1976.

[H4] D. A. Hejhal, *The Selberg Trace Formula for* $PSL(2,\mathbb{R})$, Vol. 2, Lecture Notes in Math. **1001**, Springer-Verlag, Berlin, 1983.

[H5] D. A. Hejhal, A classical approach to a well-known spectral correspondence on quaternion groups, in *Number Theory,* (edited by D. Chudnovsky, G. Chudnovsky, H. Cohn, M. Nathanson), Lecture Notes in Mathematics **1135**, Springer-Verlag, Berlin–Heidelberg–New York, 1985, 127–196.

[H6] M. N. Huxley, Scattering Matrices for Congruence Subgroups, in *Modular Forms,* (edited by R. Rankin) Ellis Horwood Ltd., Chichester, 1984, 141–156.

[IR] K. Ireland, M. Rosen, *A Classical Introduction to Modern Number Theory,* 2nd ed., Springer-Verlag, Berlin-Heidelberg-New York, 1990.

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

[JL] H. Jacquet, R. P. Langlands, *Automorphic Forms on* $GL(2)$, Lecture Notes in Math. **114**, Springer-Verlag, Berlin, 1970.

[K] S. Katok, *Fuchsian Groups,* Univ. of Chicago Press, Chicago and London, 1992.

[L1] E. Landau, *Vorlesungen über Zahlentheorie,* vols. 1-3, S.Hirzel, Leipzig, 1927.

[L2] S. Lang, *Algebraic Number Theory,* 2nd ed., Springer-Verlag, Berlin-Heidelberg-New York, 1994.

[M1] T. Miyake, On automorphic forms on $GL_2$ and Hecke operators, Ann. Math. **94** (1971), 174–189.

[M2] T. Miyake, *Modular Forms,* Springer-Verlag, New York, 1989.

[N1] T. Nagell, *Introduction to Number Theory,* John Wiley & Sons, Inc., New York, 1951.

[N2] J. Neukirch, *Algebraic Number Theory,* Springer-Verlag, Berlin, 1999.

[O1] A. Orihara, On the Eisenstein series for the principal congruence subgroups, Nagoya Math. J. **34** (1969), 129–142.

[O2] W. F. Osgood, *Lehrbuch der Funktionentheorie,* vol. 1, 5th ed, Chelsea Reprint, 1965.

[P1] Y. N. Petridis, On squares of eigenfunctions for the hyperbolic plane and a new bound on certain $L$-series, Inter. Math. Res. Notices (1995), 111–127.

[P2] Y. N. Petridis, Fourier coefficients of cusp forms, in *Number theory,* (edited by R. Gupta, K. S. Williams), CRM Proc. Lecture Notes **19**, Amer. Math. Soc., Providence, RI, 1999, 297–307.

[S1] P. Sarnak, Class numbers of indefinite binary quadratic forms, J. Number Theory **15** (1982), 229–247.

[S2] G. Shimura, On the zeta-functions of the algebraic curves uniformized by certain automorphic functions, J. Math. Soc. Japan **13** (1961), 275–331.

[S3] C. L. Siegel, Indefinite quadratische Formen und Funktionentheorie I, Math. Ann. **124** (1951), 17–54.

[S4] C. L. Siegel, Indefinite quadratische Formen und Modulfunktionen, in *Studies and Essays Presented to R. Courant on his 60th Birthday,* Interscience, 1948, 395–406.

[S5] A. Strömbergsson, *The Selberg Trace Formula for Modular Correspondences,* Licentiat thesis, Uppsala University, 1998, [http://www.math.uu.se/∼andreas/papers.html].

[S6] A. Strömbergsson, Some remarks on a spectral correspondence for Maass Waveforms, (2000)

[T1] E. C. Titchmarsh, *The Theory of Functions,* 2nd ed., Oxford University Press, 1952.

[T2] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function,* Oxford University Press, 1986.

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

[W1] A. Weil, *Basic Number Theory,* 3rd ed., Springer-Verlag, New York, 1974.

[W2] E. Weiss, *Algebraic Number Theory,* McGraw-Hill, New York, 1963.