The Eichler-Selberg Trace Formula for Level-One Hecke Operators

Alex Barron

May 20, 2013

1 Introduction

This paper explains the steps involved in the proof of the Eichler-Selberg Trace Formula for Hecke operators of level one. It is based on an appendix in Serge Lang's *Introduction to Modular Forms* written by Don Zagier, though I also draw heavily from sections of Toshitsune Miyake's *Modular Forms* and Xueli Wang's and Dingyi Pei's *Modular Forms with Integral and Half-Integral Weights*.

Section 2 summarizes the necessary background in the theory of modular forms. The material covered here is standard, so I've left out most of the details and proofs. Most of the section is based on Chapter VII of Jean-Pierre Serre's *A Course in Arithmetic*.

Section 3 covers the first major step in Zagier's terse proof of the Eichler-Selberg Formula. I've tried to explain where Zaiger's proof comes from and why such a formula should even exist, since the original proof lacks this sort of motivation for the reader.

Section 4 completes the proof of the E-S Formula. This last part requires several results from the theory of binary quadratic forms. There are several well-known number theory books that cover these facts in detail, so I assume them without proof.

2 Modular Forms and Hecke Operators

In this section we review the basic properties of modular forms and Hecke operators that we need for the rest of the paper. Most of the proofs and details are skipped over.

2.1 Modular Forms

Let $\Gamma = SL_2(\mathbb{Z})$ be the **modular group** defined

$$\Gamma \;=\; \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a, b, c, d \in \mathbb{Z},\ ad - bc = 1 \right\}.$$

Then if $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ is the upper half-plane of the complex numbers, we can let Γ act on points $z \in \mathbb{H}$ on the left: if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, then we define

$$\gamma z \ := \ \frac{az+b}{cz+d}.$$

This left-action is well-defined on $\mathbb{H}$, since

$$\Im(\gamma z) \ = \ \frac{\Im(z)}{|cz+d|^2},$$

and therefore Γ maps $\mathbb{H}$ to $\mathbb{H}$. As expected, the action of the modular group on the upper half-plane yields an equivalence relation: we say that z and w in $\mathbb{H}$ are equivalent under Γ if $z = \gamma w$ for some $\gamma \in \Gamma$.

There is an important geometrical interpretation of the set $\Gamma\backslash\mathbb{H}$ of classes or Γ-orbits of points in H. The region

$$F \ = \ \{z \in \mathbb{H} \mid \ -\frac{1}{2} \leq z < \frac{1}{2}, \ |z| > 1\}$$

contains one and only one representative from almost each class in $\Gamma\backslash\mathbb{H}$ (we have to be careful with the boundary points on the unit disc). Such a region in the complex plane is called a **fundamental domain**.

Now if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ and $f(z)$ is a meromorphic function on $\mathbb{H}$, recall that we say that f is **weakly modular of weight k** if

$$f(\gamma z) \ = \ (cz+d)^k f(z)$$

for some integer k. If we consider the elements

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \ \text{ and } \ S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

of Γ, we see that any weakly modular function $f(z)$ of weight k on $\mathbb{H}$ must satisfy the following two identities:

$$f(Tz) = f(z+1) = f(z) \ \text{ and } \ f(Sz) = f(-\frac{1}{z}) = (-z)^k f(z).$$

These results tell us that 0 is the only weakly modular function of odd weight and, most importantly, that all such functions f are periodic. A weakly modular function f therefore has a Fourier expansion

$$f(q) \ = \ \sum_{n=-\infty}^{\infty} a_n q^n, \tag{1}$$

where $q = e^{2\pi i z}$. The transformation $z \mapsto e^{2\pi i z}$ maps $\mathbb{H}$ to the unit disc $|q| < 1$ with the origin removed, and the series $f(q)$ is meromorphic on this disc.

If f extends to a meromorphic function at the point $q = 0$, we say that f is **meromorphic at ∞**. Likewise, if f extends to a holomorphic function at $q = 0$,

we say that f is **holomorphic at** ∞. This paves the way for some important definitions: if a function $f(z)$ is weakly modular of weight k, we say that f is **modular** if f is meromorphic at ∞. In this case, we let $f(\infty)$ be the value of (1) when $q = 0$. A modular function f which is holomorphic everywhere, including ∞, is called a **modular form** (since f is holomorphic, $f(\infty) = a_0$). If a modular form $g(z)$ is 0 at the cusp point ∞, then we call g a **cusp form**. We can also think of a modular function as a complex-valued function on the set Ψ of $\mathbb{C}$-lattices that obeys an analogous transformation property ("**homogeneous of degree** k"); or as a complex-valued function on isomorphism classes of elliptic curves in the case $k = 0$. These last two characterizations of modular functions are important, but we do not need to go into details for this paper.

Example 2.1. *The Eisenstein Series.*

Let $k \geq 4$ be an even integer. Then the **Eisenstein Series**

$$G_k(z) \ := \sum_{\substack{(n,m) \\ (n,m)\neq(0,0)}} \frac{1}{(nz+m)^k}$$

is a modular form of weight k. $G_k(z)$ has the Fourier expansion

$$G_k(z) \ = \ 2\zeta(k) + A_k \sum_{n=0}^{\infty} \sigma_{k-1}(n) q^n,$$

where A_k is a constant that depends on k, ζ is the Riemann zeta function, and $\sigma_{k-1}(n)$ is the **sum of divisors function**

$$\sigma_{k-1}(n) \ := \ \sum_{d|n} d^{k-1}$$

from number theory. Note that $0 \neq \zeta(k) < \infty$ for even k, and that $G_k(\infty) = 2\zeta(k)$. Therefore G_k is not a cusp form.

The Eisenstein series gives us one example of a much more general trend in the theory of modular forms: the Fourier coefficients of these functions often carry interesting information for number theorists (like the value of the sum of divisors function). It is therefore useful to develop methods and algorithms that help us compute these coefficients. The central theorem of the paper gives us one such method, as we will see below.

Moving on for now, let M_k be the set of all modular forms of weight k, and let $S_k \subset M_k$ be the set of all cusp forms of weight k. Then we can give M_k the structure of a $\mathbb{C}$-vector space using point-wise addition. It turns out that M_k is finite-dimensional, which is rare for spaces of functions. S_k is a subspace of M_k, and in fact forms a "big" piece of the space. If $\varphi : M_k \to \mathbb{C}$ is the linear functional defined $\varphi(f) = f(\infty)$, then S_k is the kernel of φ, and therefore if $S_k \subsetneq M_k$ then

$$M_k/S_k \ \cong \mathbb{C}.$$

We saw in the above example that when $k \geq 4$, the Eisenstein series $G_k(z)$ is in M_k and not in S_k. As a consequence we have the decomposition

$$M_k \;=\; S_k \oplus \mathbb{C}G_k$$

for $k \geq 4$.

Since we know a lot about the Eisenstein series, we can focus most of our attention on the space S_k of cusp forms. It turns out that S_k is an inner-product space: for $f, g \in S_k$, we let

$$\langle f, g \rangle \;:=\; \int_F f(z)\overline{g(z)}y^k \; \frac{dxdy}{y^2},$$

where F is a fundamental domain for Γ in $\mathbb{H}$ and $z = x+iy$ (the reader can check that the differential $\dfrac{dxdy}{y^2}$ is invariant under the action of Γ). This operation is called the **Petersson inner product**. The value of $\langle f, g \rangle$ does not depend on the choice of domain F.

2.2 Hecke Operators

Let $k \geq 4$ be an even integer. A **Hecke operator** is a special linear transformation that acts on the space M_k of modular forms. If m is a positive integer and $F(\Lambda)$ is a homogeneous function of degree k on the set Ψ of $\mathbb{C}$-lattices, we can map F to another such function using the transformation T_m defined

$$T_m F(\Lambda) \;:=\; \sum_{[\Lambda \,:\, \Lambda'] = m} F(\Lambda'),$$

where the sum is over sub-lattices of Λ of index m. We can interpret T_m as an operator $T_k(m)$ on the space M_k using the correspondence between lattice functions and modular forms that we mentioned above. This requires some work, which we omit.

If f is a modular form of weight k, it turns out that we can express the action of $T_k(m)$ on f as the sum

$$T_k(m)f(z) \;=\; m^{k-1} \sum_{\substack{a \geq 1, ad = m \\ 0 \leq b \leq d}} d^{-k} f\left(\frac{az+b}{d}\right),$$

where a, b, and d are integers. $T_k(m)f$ is an element of M_k, and if $f \in S_k$ then $T_k(m)f$ is a cusp form of weight k. The Hecke operators have a number of powerful properties which we list below:

1. Multiplicativity: if $\gcd(n, m) = 1$, then

$$T_k(mn) = T_k(m)T_k(n).$$

2. $T_k(m)$ and $T_k(n)$ commute for all m and n.

3. If p is prime and $n \geq 1$, then

$$T_k(p^n)T_k(p) = T_k(p^{n+1}) + p^{1-k}T_k(p^{n-1}).$$

4. $T_k(m)$ is hermitian for each integer $m \geq 1$.

In addition to all of this, we have the following result about the eigenvalues of the Hecke operators:

Theorem 2.1. *Let*

$$f(z) \;=\; \sum_{n=0}^{\infty} c(n)q^n$$

be a modular form of weight k (where k is an even integer), and assume that f is an eigenfunction of all the $T_k(m)$, i.e. that there exists a complex number $\lambda(m)$ such that

$$T_k(m)f = \lambda(m)f$$

for all integers $m \geq 1$. Then

1. *The coefficient $c(1)$ of q in f nonzero.*

2. *If f is normalized by the condition $c(1) = 1$, then*

$$c(m) \;=\; \lambda(m)$$

for all $m > 1$.

This theorem tells us that the eigenvalues of Hecke operators are Fourier coefficients of modular forms. We saw in Example 2.1 that these coefficients often contain interesting information for number theorists, and this is one of the main reasons why mathematicians study Hecke operators. The Eisenstein series $G_k(z)$ is one example of an eigenfunction of the $T_k(m)$, with eigenvalues $\lambda(m) = \sigma_{k-1}(m)$.

3 Preliminary Trace Formula

As we've just seen above, the eigenvalues of Hecke operators carry a lot of interesting arithmetical information. In practice, however, it is often difficult to determine and compute these numbers directly. We need to use alternative methods, which is why mathematicians are so interested in studying the trace of Hecke operators; if we can't get at the eigenvalues directly, we can at least get at their sum.

In this section we derive a preliminary formula for the trace of Hecke operators acting on the space S_k of cusp forms of weight k. The formula is "preliminary" in the sense that it doesn't actually tell us how to compute the trace.

Instead, it is more of a means to set up the problem for the analysis that follows in Section 4.

The proof is mostly self-contained, but the first few steps draw heavily from the theory of function spaces. In 3.1 we summarize the necessary background in analysis, paving the way for the derivation of the initial trace formula in 3.2. The goal of 3.1 is to give the reader a general sense of where the preliminary trace formula comes from, and for this reason we list many powerful results without proof. The interested (or worried) reader can find all of these proofs in [Miyake].

3.1 Spaces of Functions on the Upper Half-Plane

Let k be a non-negative real number, and suppose p is an element of the interval $[1, \infty]$ of real numbers adjoined with ∞. We call the number p an **exponent**. If $f(z)$ is a function on the upper half-plane $\mathbb{H}$, we define a norm

$$\|f\|_p \;=\; \begin{cases} \left(\int\limits_{\mathbb{H}} |f(z)\Im(z)^{k/2}|^p \; \dfrac{dxdy}{y^2} \right)^{1/p} & \text{if } \; 1 \leq p < \infty \\ \\ \operatorname{ess.sup}\limits_{z \in \mathbb{H}} \; |f(z)\Im(z)^{k/2}| & \text{if } \; p = \infty \end{cases} . \tag{2}$$

This is a slight variation of the standard "p-norm."

We let $L^p_k(\mathbb{H})$ denote the set of all measurable functions f on $\mathbb{H}$ satisfying $\|f\|_p < \infty$. In the case $k = 0$, the set $L^p_0(\mathbb{H})$ is just the usual L^p-space of functions on $\mathbb{H}$. In particular, we know that $L^p_0(\mathbb{H})$ is a Banach space (a complete normed vector space). Since $L^p_0(\mathbb{H})$ and $L^p_k(\mathbb{H})$ are isomorphic as normed spaces by the map $f(z) \mapsto f(z)\Im(z)^{k/2}$, it follows that $L^p_k(\mathbb{H})$ is also a Banach space.

If q is another exponent such that

$$\frac{1}{p} + \frac{1}{q} \;=\; 1,$$

then we say that p and q are **conjugate** exponents (if $p = \infty$ then we let $\frac{1}{p} = 0$, so $q = 1$). If we have two such conjugate exponents p and q, then if $f \in L^p_k(\mathbb{H})$ and $g \in L^q_k(\mathbb{H})$, we define a product

$$(f, g)_{\mathbb{H}} \;=\; \int\limits_{\mathbb{H}} f(z)\overline{g(z)}\Im(z)^k \; \frac{dxdy}{y^2}. \tag{3}$$

With this pairing, we can identify an element $g \in L^q_k(\mathbb{H})$ with a continuous linear functional in the dual space of $L^p_k(\mathbb{H})$. Conversely, it can be shown that if $p \neq \infty$, then for any functional α on $L^p_k(\mathbb{H})$ there is a $g_\alpha \in L^q_k(\mathbb{H})$ such that $(f, g_\alpha) = \alpha(f)$. So if $p \neq \infty$ and p and q are conjugates, then $L^q_k(\mathbb{H})$ is the dual space of $L^p_k(\mathbb{H})$ (and vice versa). Additionally, in the case $p = q = 2$, our product $(\ ,\ )_{\mathbb{H}}$ is an inner product that makes $L^2_k(\mathbb{H})$ a Hilbert space.

Now let $H_k^p(\mathbb{H})$ denote the set of holomorphic functions in $L_k^p(\mathbb{H})$. It turns out that $H_k^p(\mathbb{H})$ is a closed subspace of $L_p^k(\mathbb{H})$, which tells us that the holomorphic functions on $\mathbb{H}$ that behave well with respect to the norm $\|\cdot\|_p$ form a Banach space. In order to see why this is true, we need a few general theorems from complex analysis:

Theorem 3.1. *Let k and p be real numbers. If U is an open subset of $\mathbb{H}$ and $f(z)$ is holomorphic on U, then for every $z_0 \in U$ there is a constant C_{z_0} such that*

$$|f(z_0)| \ \leq \ C_{z_0} \left(\int\limits_U |f(z)\Im(z)^{k/2}| \ \frac{dxdy}{y^2} \right)^{1/p}.$$

Theorem 3.2. *Let k and p be real numbers. Assume that $f(z)$ is holomorphic on $\mathbb{H}$ and that*

$$\int\limits_{\mathbb{H}} |f(z)\Im(z)^{k/2}|^p \ \frac{dxdy}{y^2} < \infty.$$

Then for any real numbers a and b such that $0 < a < b$, we have

$$\lim_{|x|\to\infty} |f(z)| = 0$$

uniformly with respect to y on the interval $[a, b]$.

Theorem 3.3. *Let k and p be real numbers, and suppose U is a subdomain of $\mathbb{H}$. For a holomorphic function $f(z)$ on U, we put*

$$\|f\|_U \ = \ \left(\int\limits_U |f(z)\Im(z)^{k/2}|^p \ \frac{dxdy}{y^2} \right)^{1/p}.$$

Let M be a compact subset of U. If a sequence $\{f_n(z)\}$ of holomorphic functions on U is Cauchy with respect to the norm $\|\cdot\|_U$, then $\{f_n(z)\}$ converges uniformly on M. As a consequence, the limit function of $\{f_n(z)\}$ is holomorphic on U.

The proof of Theorem 3.2 requires the bound from Theorem 3.1, and Theorem 3.3 is an application of Theorem 3.2. The reader can find detailed proofs in [Miyake]. Shifting our focus back to the space $H_k^p(\mathbb{H})$, we immediately have the following:

Theorem 3.4. *The space $H_k^p(\mathbb{H})$ is a closed subspace of $L_k^p(\mathbb{H})$.*

Proof. Suppose that $f(z)$ is the limit of a sequence of functions $\{f_n(z)\}$ in $H_k^p(\mathbb{H})$. Using Theorem 3.3, we see that $f(z)$ is also holomorphic. Therefore $f(z) \in H_k^p(\mathbb{H})$. □

One consequence of this theorem is that $H_k^2(\mathbb{H})$ is a Hilbert space with the inner product $(\ ,\ )_{\mathbb{H}}$ from above. It turns out that we can associate $H_k^2(\mathbb{H})$ with a special function called the "kernel" of the Hilbert space. These "kernel"

functions carry information about every element of the space, and for this reason they are incredibly useful. Here's the formal definition:

Suppose that H is a Hilbert space of complex-valued functions on a set X, and let $(\ ,\ )_H$ be the inner-product of H. We call a function $K(x,y)$ on $X \times X$ a **kernel function** of H if, for any $y \in X$, it satisfies the following two conditions:

1. $K(x,y)$ is an element of H as a function of x.
2. If $f(x) \in H$, then $f(y) = (f(x), K(x,y))_H$. Here the inner product is taken with respect to the variable x.

We should point out that a Hilbert space H does not necessarily have a kernel function. Luckily for us, it isn't too hard to show that $H_k^2(\mathbb{H})$ does.

Proposition 3.1. *The space $H_k^2(\mathbb{H})$ has a kernel function.*

Proof. Since 2 is its own conjugate, we can use the inner product to interpret $H_k^2(\mathbb{H})$ as its own dual space. If we pick an $f \in H_k^2(\mathbb{H})$ and $z_0 \in \mathbb{H}$, it therefore follows that there is a $g_{z_0} \in H_k^2(\mathbb{H})$ such that

$$f(z_0) \ = (f, g_{z_o})_{\mathbb{H}}.$$

We set $K_k(z_1, z_2) = g_{z_2}(z_1)$. Then $K_k \in H_k^2(\mathbb{H})$ for fixed z_2, and if $f \in H_k^2(\mathbb{H})$ then

$$(f(z_1), K_k(z_1, z_2))_{\mathbb{H}} \ = \ (f, g_{z_2})_{\mathbb{H}} \ = \ f(z_2).$$

So $K_k(z_1, z_2)$ is the kernel function we're looking for. (Note that this proof doesn't work for the usual L^2 space, since, for example, "$f(z_0)$" doesn't make sense for $f \in L^2$). □

This existence proof is nice, but it doesn't give us a concrete formula for the kernel $K_k(z_1, z_2)$. It is possible to find such a formula, but the derivation requires an extensive amount of set-up. For the sake of space and cohesion, we refer the reader to [Miyake] for all of the work. Miyake proves that the kernel of $H_k^2(\mathbb{H})$ is the function

$$K_k(z_1, z_2) \ = \ \frac{k-1}{4\pi} \left(\frac{z_1 - \overline{z_2}}{2i} \right)^{-k},$$

where $z_1, z_2 \in \mathbb{H}$.

The theory behind this result is crucial to understanding the derivation of the trace formula in the next section, as we will soon see. The goal now is to adapt everything we've done so far to the space S_k of cusp forms of weight k. For our purposes this only requires a small adjustment, but we should point out that this is the launching point for a much more general application of L^p theory to spaces of modular forms (see [Miyake]).

Let $\Gamma = SL_2(\mathbb{Z})$, and suppose $f(z)$ is a modular function of weight k. If p is an exponent, we define the norm $\|f\|_{\Gamma,p}$ in almost the same way as $\|f\|_p$, except we integrate over a fundamental domain F of Γ in $\mathbb{H}$. Since $|f(z)\Im(z)^{k/2}|$ is

invariant under the action of Γ, the norm $\|f\|_{\Gamma,p}$ is well-defined. We let $L_k^p(\Gamma)$ denote the set of measurable, modular functions $f(z)$ of weight k for which $\|f\|_{\Gamma,p} < \infty$. As before, we let $H_k^p(\Gamma)$ be the subset of $L_k^p(\Gamma)$ containing all the holomorphic elements. $L_k^p(\Gamma)$ is a Banach space and $H_k^p(\Gamma)$ is a closed subspace, by the same arguments as above.

In particular, we know that $L_k^2(\Gamma)$ is a Hilbert space with the inner product

$$(f,g)_\Gamma \;=\; \int\limits_F f(z)\overline{g(z)}\Im(z)^k \;\frac{dxdy}{y^2}. \tag{4}$$

Since the volume of F is finite, $L_k^\infty(\Gamma) \subset L_k^p(\Gamma)$ and $H_k^\infty(\Gamma) \subset H_k^p(\Gamma)$ for $1 \leq p < \infty$. Since cusp forms decrease rapidly to 0 as $\Im(z) \to \infty$, we can show that $S_k \subset H_k^\infty(\Gamma)$. It also turns out that $H_k^\infty(\Gamma) \subset S_k$ and hence that $H_k^\infty(\Gamma) = S_k$, but the full proof of this result will take us off track. The interested reader should check Theorems 2.1.4 and 2.1.5 in [Miyake] for the details.

As a consequence, the cusp forms of weight k are embedded in $L_k^p(\Gamma)$ for any exponent p. In the case $p = 2$, the restriction of the inner product of $L_k^2(\Gamma)$ to $H_k^\infty(\Gamma)$ coincides with the Petersson inner product (up to a constant multiple). We also have the following:

Theorem 3.5. *$H_k^2(\Gamma) = H_k^\infty(\Gamma)$.*

Proof. Given the discussion above, we want to show that if $f \in H_k^2(\Gamma)$ then f is zero at ∞. It will follow that f is a cusp form, and therefore that $f \in S_k = H_k^\infty(\Gamma)$.

The intuitive idea is that

$$\int\limits_1^\infty \int\limits_{-1/2}^{1/2} |f(z)|^2 y^{k-2}\; dxdy \;<\; \infty,$$

and therefore $|f(z)|^2$ must decrease rapidly to 0 as $y = \Im(z) \to \infty$. For the details, see 6.3.1 in [Miyake]. □

This result tells us that $S_k = H_k^2(\Gamma)$. We saw above that the space $H_k^2(\mathbb{H})$ has a kernel function, so it is reasonable to ask if S_k does as well. In the next section we show that S_k does have such a function. We then use this kernel to derive our preliminary formula for the trace of Hecke operators.

3.2 The Kernel Function of S_k

Let $k \geq 4$ be an even integer. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ and $z = x + iy \in \mathbb{H}$, we let $J(\gamma, z) = cz + d$. For z and w in the upper half-pane, we define a function

$$h_1(z,w) \;=\; N_k^{-1} \sum_{\gamma \in \Gamma} K_k(\gamma z, w) J(\gamma, z)^{-k}, \tag{5}$$

where $K_k(z,w)$ is the kernel of $H_k^2(\mathbb{H})$ and $N_k = \frac{2^{k-2}(k-1)}{(-1)^{k/2}\pi}$. We will ultimately show that $N_k h_1(z,w)$ is the kernel function of S_k, but this requires some work. We first show that h_1 is modular of weight k in each variable, and then prove that it is a cusp form in each variable.

Recall from the previous section that

$$K_k(z,w) \;=\; \frac{k-1}{4\pi}\left(\frac{z-\overline{w}}{2i}\right).$$

We can therefore write

$$h_1(z,-\overline{w}) \;=\; \sum_{\gamma\in\Gamma} J(\gamma,z)^{-k}(w+\gamma z)^k,$$

noting that $-\overline{w}\in\mathbb{H}$ since $w\in\mathbb{H}$. If we let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ run over Γ, then we can rewrite this sum as

$$h_1(z,-\overline{w}) \;=\; \sum_{ad-bc=1} (cz+d)^{-k}\left(w+\frac{az+b}{cz+d}\right)^{-k}. \tag{6}$$

Multiplying through by $(cz+d)^{-k}$ gives

$$h_1(z,-\overline{w}) \;=\; \sum_{ad-bc=1} (czw+dw+az+b)^{-k}, \tag{7}$$

which tells us that $h_1(z,-\overline{w})$ is symmetric in z and w (this change of variables amounts to switching a and d in the sum (7)). Therefore, if we can show that h_1 is modular in z, it will follow that it is modular in both variables. This is a straightforward argument:

If $A\in\Gamma$ and $B\in\Gamma$, then it is easy to verify that

$$(AB)z = A(Bz)$$

and

$$J(AB,z) = J(A,Bz)J(B,z).$$

We then have

$$\begin{aligned} h_1(Bz,-\overline{w}) \;&=\; \sum_{A\in\Gamma} J(A,Bz)^{-k}(w+ABz)^{-k} \\ &= \sum_{A\in\Gamma}\left(\frac{J(AB,z)}{J(B,z)}\right)^{-k}(w+ABz)^{-k} \\ &= J(B,z)^k \sum_{A\in\Gamma} J(AB,z)^{-k}(w+ABz)^{-k} \\ &=\; J(B,z)^k h_1(z,-\overline{w}), \end{aligned}$$

proving modularity in z and hence in $-\overline{w}$. Of course, this result doesn't tell us much if the function h_1 diverges in $\mathbb{H}$. The next proposition resolves this issue, and allows us to conclude the proof that $h_1(z,w)$ is a cusp form of weight k in each variable.

Proposition 3.2. *The function $h_1(z,w)$ is absolutely and uniformly convergent on compact subsets of $\mathbb{H}\times\mathbb{H}$. As a consequence, $h_1(z,w) \in H_k^2(\mathbb{H})$ as a function of both z and w.*

Proof. Let F be a fundamental domain for $\Gamma \in \mathbb{H}$, and set $z = x+iy$ as always. For fixed w, we have

$$\int_F |h_1(z,w)|y^{k/2}\ \frac{dxdy}{y^2} \ \leq\ \int_F \sum_{\gamma\in\Gamma} |K_k(\gamma z,w)J(\gamma,z)^{-k}|y^{k/2}\ \frac{dxdy}{y^2}$$

$$= \sum_{\gamma\in\Gamma}\int_F |K_k(\gamma z,w)J(\gamma,z)^{-k}|y^{k/2}\ \frac{dxdy}{y^2}.$$

Now if $\gamma\in\Gamma$, then the reader can check that

$$K_k(\gamma z,w)J(\gamma,z)^{-k} \ =\ K_k(z,\gamma^{-1}w)\overline{J(\gamma^{-1},w)^{-k}}$$

and

$$K_k(\gamma z,\gamma w) \ =\ K_k(z,w)J(\gamma,z)^k\overline{J(\gamma,w)^k}.$$

Miyake gives a proof of these identities in Section 6.1 of [Miyake] that avoids a messy computation. Applying these two results, we see that

$$\int_F |h_1(z,w)|y^{k/2}\ \frac{dxdy}{y^2} \ \leq\ \sum_{\gamma\in\Gamma}\int_F |K_k(z,\gamma^{-1}w)\overline{J(\gamma^{-1},w)^{-k}}|y^{k/2}\ \frac{dxdy}{y^2}$$

$$= \sum_{\gamma\in\Gamma}\int_{\gamma F} |K_k(\gamma^{-1}z,\gamma^{-1}w)\overline{J(\gamma^{-1},w)^{-k}}|\ \frac{y^{k/2}}{|J(\gamma^{-1},z)|^k}\ \frac{dxdy}{y^2}$$

$$= \sum_{\gamma\in\Gamma}\int_{\gamma F} |K_k(z,w)J(\gamma^{-1},z)^k|\frac{y^{k/2}}{|J(\gamma^{-1},z)|^k}\ \frac{dxdy}{y^2}$$

$$= \sum_{\gamma\in\Gamma}\int_{\gamma F} |K_k(z,w)|y^{k/2}\ \frac{dxdy}{y^2}$$

$$= \int_{\mathbb{H}} |K_k(z,w)|y^{k/2}\ \frac{dxdy}{y^2} \ <\ \infty$$

since $K_k(z,w) \in H_k^2(\mathbb{H})$. As a consequence, $h_1(z,w)$ must converge absolutely. Moreover, suppose U is a compact subset of the fundamental domain F. Then since

$$\int_U |h_1(z,w)| y^{k/2} \, \frac{dxdy}{y^2} \; < \; \infty,$$

Theorem 3.3 implies that h_1 converges uniformly on U as a function of z (using the case $p = 1$). But h_1 is symmetric in each variable, and therefore the same result follows for h_1 as a function of w. Since h_1 is modular in z and w, it follows that $h_1(z,w)$ converges uniformly on compact subsets of $\mathbb{H} \times \mathbb{H}$. As a consequence, $h_1(z,w)$ is holomorphic in $\mathbb{H}$ as a function of z and w. □

All that's left to show is that h_1 is holomorphic and zero at the cusp point ∞. It is intuitively obvious from (6) that h_1 equals 0 at the points $z = \infty$ and $w = \infty$, and Theorem 3.5 along with Proposition 3.2 allow us to conclude that h_1 is holomorphic at ∞. This argument is a little sketchy, largely because it draws from an enormous amount of background material that we only briefly summarized in the last section (as does the proof of Proposition 3.2). The missing details are provided in Sections 2.6 and 6.3 in [Miyake]. For now, we will conclude that $h_1(z,w) \in S_k$ as a function of both z and w.

To prove that $h_1(z,w)$ is the kernel of S_k (up to a constant), we will consider the action of the m-th Hecke operator $T_k(m)$ on $S_k = H_k^2(\Gamma)$ and express it in terms of a product $*$ with a generalization of the function h_1. Since $T_k(1)$ is just the identity transformation, we can use this result to argue that h_1 is the kernel function of S_k. In the process, we will develop enough machinery to derive a preliminary formula for the trace of $T_k(m)$.

For positive integers m, we let $\mathbb{M}^m$ denote the set of all integer matrices of determinant m. We also let $T(m) = T_k(m)$ be the m-th Hecke operator acting on S_k. Since $h_1 \in S_k$, the function $(T(m)h_1)(z)$ is a cusp form for each m (we let $T(m)$ act with respect to the first variable z of $h_1(z,w)$, but this choice is arbitrary since h_1 is symmetric). Now if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ ranges over $\mathbb{M}^m$, we define a new function

$$h_m(z,w) \; := \sum_{ad-bc=m} (cz+d)^{-k} \left(w + \frac{az+b}{cz+d} \right)^{-k}. \tag{8}$$

This is a natural extension of the definition (6) of h_1.

Proposition 3.3. *For any positive integer m,*

$$m^{k-1} h_m \; = \; T(m) h_1,$$

where $T(m)$ operates with respect to the first variable z on the right-hand side. Therefore h_m is a cusp form of weight k in each variable.

Proof. Suppose m is a positive integer and write

$$m^{k-1} h_m(z,w) \; = \; m^{k-1} \sum_{M \in \mathbb{M}^m} J(M,z)^{-k} (w + Mz)^{-k}. \tag{9}$$

Let $\Gamma\backslash\mathbb{M}^m$ be the set of congruence classes of $\mathbb{M}^m$ under the left action of Γ. Then we can break right side of (9) into

$$m^{k-1} \sum_{\Lambda\in\Gamma\backslash\mathbb{M}^m} \sum_{A\in\Lambda} J(A,z)^{-k}(w+Az)^{-k}.$$

Now recall that the integer matrices $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ with $a>0$, $ad=m$, and $0\le b<d$ form a complete set of representatives of classes in $\Gamma\backslash\mathbb{M}^m$ (see [Serre]). As a consequence, we have

$$m^{k-1}h_m(z,w) \;=\; m^{k-1} \sum_{\substack{a>0,ad=m \\ 0\le b<d}} \sum_{BM} J(BM,z)^{-k}(w+BMz)^{-k},$$

where M is the matrix $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ and B runs over Γ. Using the identities from the proof of modularity of h_1,

$$m^{k-1}h_m(z,w) \;=\; m^{k-1} \sum_{\substack{a>0,ad=m \\ 0\le b<d}} \sum_{BM} J(B,Mz)^{-k}J(M,z)(w+B(Mz))^{-k}.$$

Since $J(M,z)^{-k}=d^{-k}$ and

$$\sum_{BM} J(B,Mz)^{-k}(w+B(Mz))^{-k} \;=\; h_1(Mz,-\overline{w}),$$

we find that

$$m^{k-1}h_m(z,w) \;=\; m^{k-1} \sum_{\substack{a>0,ad=m \\ 0\le b<d}} d^{-k}h_1(Mz,-\overline{w})$$

$$= T(m)h_1,$$

where $T(m)$ acts on h_1 with respect to z. Therefore $h_m\in S_k$ as a function of z. Since h_m is symmetric in z and w (for the same reasons as h_1), it follows that $h_m\in S_k$ as a function of w as well. □

The next step is to extend the Petersson inner product on $S_k=H_k^2(\Gamma)$ to a function of a variable w in the upper half-plane. Suppose $h(z,w)$ is a function of two variables on $\mathbb{H}$, and that h is an element of S_k in each variable. Then if $f\in S_k$, we define $f*h$ as a function of w by setting

$$f*h(w) \;:=\; \int_F f(z)\overline{h(z,\overline{-w})}y^k\,\frac{dxdy}{y^2}, \tag{10}$$

where F is a fundamental domain for Γ in $\mathbb{H}$, $x=\Re(z)$, and $y=\Im(z)$ as before. If we fix w and treat h as a function of z, then this operation is merely the Petersson product of f and h, as claimed.

The next theorem is the key result of the section. It tells us that if f is a cusp form of weight k, then we can identify the product $f * h_m(w)$ with the action of the Hecke operator $T(m)$ on f, up to a constant that depends only on k and m. It follows that $N_k h_1$ is the kernel function of S_k, and also that h_m is a kind of "kernel function" for the operator $T(m)$.

Theorem 3.6. *If m is a positive integer and $f \in S_k$, then*

$$f * h_m(w) \;=\; C_k m^{-k+1}(T(m)f)(w),$$

where C_k is a constant that depends only on k. In particular, h_1 is the kernel function of S_k (up to a constant).

Proof. Suppose f is a cusp form of weight k. The trick is to first consider the case where $m = 1$: we know from the above proposition that $T(m)h_1 = m^{k-1}h_m$, and we can use this result to write $f * h_m$ in terms of $T(m)f * h_1$.

In this case $T(1)$ is just the identity transformation, so we want to show that $f * h_1(w) = C_k f(w)$. Let us try, then, to simplify and compute the integral

$$f * h_1(w) \;=\; \int_F f(z)\overline{h_1(z, \overline{-w})}y^k \; \frac{dxdy}{y^2}.$$

We begin with a few identities. For any $z = x + iy$ and $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma = SL_2(\mathbb{Z})$, we have

$$f(Mz) \;=\; (cz + d)^k f(z)$$

and

$$\Im(Mz)^k \;=\; \frac{\Im(z)^k}{|cz + d|^{2k}},$$

so that

$$f(Mz)\Im(Mz)^k \;=\; (c\overline{z} + d)^{-k} f(z)y^k. \tag{11}$$

Since $\overline{h_m(z, w)} = h_m(\overline{z}, \overline{w})$, (11) and (6) imply that

$$\begin{aligned} f(z)\overline{h_1(z, w)}y^k \;&=\; \sum_{M\in\Gamma} (c\overline{z} + d)^{-k} f(z)y^k(\overline{w} + M\overline{z})^{-k} \\ &=\; \sum_{M\in\Gamma} (\overline{w} + M\overline{z})^{-k} f(Mz)\Im(Mz)^k. \end{aligned}$$

Using this formula, we can rewrite the product of f and h_1 as

$$f * h_1(w) \;=\; \int_F \sum_{M\in\Gamma} (-w + M\overline{z})^{-k} f(Mz)\Im(Mz)^k \; \frac{dxdy}{y^2}.$$

Since h_1 converges absolutely and uniformly, we can switch the order of integration and summation. But our integral is also invariant under the choice of domain F, so a change of variables yields

$$f * h_1(w) \;=\; \sum_{M\in\Gamma} \int_{MF} (-w + \overline{z})^{-k} f(z)\Im(z)^k \; \frac{dxdy}{y^2}.$$

Recall from Section 2 that the upper half-plane of $\mathbb{C}$ is equal to the union of all transforms of F by elements of Γ. This union is disjoint, except for a set of boundary points of measure zero in $\mathbb{C}$. Therefore $f * h_1(w)$ is actually an integral over $\mathbb{H}$, so we have

$$f * h_1(w) \;=\; 2\int_0^\infty \int_{-\infty}^\infty (x - iy - w)^{-k} f(x + iy) y^k \; \frac{dxdy}{y^2}, \tag{12}$$

where the factor of 2 comes from the fact that $MF = (-M)F$ for all $M \in \Gamma$.

We've reduced the product of f and h_1 to a more manageable form, and the rest of the proof is largely computational. Pick a real number $R > 0$, and let C be an upper-semicircle of radius R centered above the origin in $\mathbb{C}$ and oriented counter-clockwise. Then we can break the integral

$$\oint_C \frac{f(z)}{(z - 2iy - w)^k} \; dz$$

into the sum

$$\int_{-R}^{R} \frac{f(x + iy)}{(x - iy - w)^k} \; dx \;+\; \int_{\alpha_R} \frac{f(z)}{(z - 2iy - w)^k} \; dz,$$

where α_R is the arc of the semicircle. Since $f(z)$ is a cusp form, it is holomorphic in $\mathbb{H}$ and rapidly approaches zero as $\Im(z) \to \infty$. Therefore, using Cauchy's inequality we find that the integral over α_R approaches zero as $R \to \infty$. But the integral over C remains constant as R grows (since the resultant semicircle can always be continuously deformed into C), so taking the limit as R approaches ∞, we have

$$\int_{-\infty}^{\infty} \frac{f(x + iy)}{(x - iy - w)^k} \; dx \;=\; \oint_C \frac{f(z)}{(z - 2iy - w)^k} \; dz.$$

By Cauchy's formula, it follows that

$$\int_{-\infty}^{\infty} f(x + iy)(x - iy - w)^{-k} \; dx \;=\; \frac{2\pi i}{(k-1)!} f^{(k-1)}(2iy + w),$$

which with (12) implies that

$$f * h_1(w) \;=\; \frac{4\pi i}{(k-1)!} \int_0^\infty f^{(k-1)}(2yi + w) y^{k-2} \; dy.$$

We now repeatedly use integration by parts to compute this integral. In the first iteration, we write $f * h_1(w)$ as

$$\frac{-4\pi i}{(k-1)!} \left(\left[\frac{y^{k-2}}{2i} f^{(k-2)}(2yi + w) \right]_0^\infty \;+\; \frac{(k-2)}{2i} \int_0^\infty y^{k-3} f^{(k-2)}(2yi + w) \; dy \right).$$

But f is a cusp form and hence zero at infinity, so the first term of the sum is zero. This fact remains true as we continue to integrate by parts, and we ultimately obtain

$$f * h_1(w) \;=\; \frac{4\pi i}{(k-1)!}\left(\frac{-1}{2i}\right)^{k-2}(k-2)!\frac{-1}{2i}f(w).$$

If we set

$$\begin{aligned} C_k \;&=\; \frac{4\pi i}{(k-1)!}\left(\frac{-1}{2i}\right)^{k-2}(k-2)!\frac{-1}{2i} \\ &=\; \frac{(-1)^{k/2}\pi}{2^{k-3}(k-1)}, \end{aligned}$$

then

$$f * h_1(w) \;=\; C_k f(w). \tag{13}$$

Since $T(1)$ is the identity transformation, this completes the case $m = 1$. It also shows that $C_k^{-1}h_1$ is the kernel of S_k.

The rest of the proof is immediate. Using (13), Proposition 3.3, and the fact that $T(m)f \in S_k$, we see that

$$\begin{aligned} f * h_m(w) \;&=\; f * (m^{-k+1}T(m)h_1)(w) \\ &=\; m^{-k+1} f * (T(m)h_1)(w) \\ &= m^{-k+1}(T(m)f) * h_1(w) \\ &= C_k m^{-k+1}(T(m)f)(w), \end{aligned}$$

as claimed. We've also used the property that $T(m)$ is hermitian, interpreting $*$ as an inner product with respect to z. □

We now use Theorem 3.6, along with some linear-algebraic properties of S_k and Hecke operators, to derive a formula for the trace of $T(m)$. Recall from above that $S_k = H_k^2(\Gamma)$ is a finite-dimensional inner product space, and that the Hecke operator $T(m)$ is hermitian for all positive integers m. For each Hecke operator, we can therefore construct a basis of eigenforms $\{f_1, f_2, ..., f_r\}$ which are orthogonal with respect to the Petersson inner product (by the Spectral Theorem). Additionally, we can assume that the f_j are normalized in the sense discussed in Theorem 2.1, that is, if

$$f_j(z) = \sum_{m=1}^{\infty} a_n^{(j)} q^n,$$

with $q = e^{2\pi iz}$ as before, then $a_1^{(j)} = 1$. We saw earlier that if an eigenform f_j is normalized, then $T(m)f_j = a_m^{(j)} f_j$. In words, the eigenvalue for f_j is its m-th Fourier coefficient. Moreover, we know that each eigenvalue $a_m^{(j)}$ is a real number, since $T(m)$ is a hermitian operator.

In particular, if we have an eigenform f_l of $T(m)$ with eigenvalue $\lambda \in \mathbb{R}$, then Theorem 3.6 tells us that $f_l * h_m(w) = C_k m^{-k+1} \lambda f_l(w)$. This fact allows us to rewrite the function $h_m(z,w)$ in a nice way with respect to the basis of eigenforms $\{f_1, f_2, ..., f_r\}$, as we show in the following lemma. The trace formula we're after will then follow easily.

Lemma 3.1. *If $\{f_1, f_2, ..., f_r\}$ is the basis of eigenforms described above, then*

$$C_k^{-1} m^{k-1} h_m(z,w) = \sum_{l=1}^{r} \frac{a_m^{(l)}}{\langle f_l, f_l \rangle} f_l(z) f_l(w),$$

where $\langle \cdot, \cdot \rangle$ is the Petersson inner product and $a_m^{(l)}$ is the eigenvalue of f_l.

Proof. The goal is to write

$$h_m(z,w) = \sum_{i,j=1}^{r} a_{ij} f_i(z) f_j(w) \tag{14}$$

for $a_{ij} \in \mathbb{C}$, and then use Theorem 3.6 to compute $f_j * h_m(w)$ for each j.

In order to see that the double sum in (14) is valid, fix w and think of h_m as a function of z. Then

$$h_m(z,w) = \sum_{i=1}^{r} a_i(w) f_i(z),$$

where the constant terms $a_i(w) \in \mathbb{C}$ depend on w. We claim that if we let w vary over $\mathbb{H}$, then $a_i(w)$ is a cusp form for each i. Since h_m is a cusp form, $a_i(w)$ must be holomorphic on $\mathbb{H}$ and zero at ∞, so we only need to show that $a_i(w)$ is modular of weight k. This follows easily, since if $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, then

$$\begin{aligned} h_m(z, Bw) &= \sum_{i=1}^{r} a_i(Bw) f_i(z) \\ &= (cw+d)^k h_m(z,w) \\ &= \sum_{i=1}^{r} (cw+d)^k a_i(w) f_i(z) \end{aligned}$$

because h_m is a cusp form in w. Therefore $a_i(Bw) = (cw+d)^k a_i(w)$ by the independence of the f_i, so $a_i(w)$ is a cusp form as claimed. As a consequence, we can expand the $a_i(w)$ as a linear combination of $f_1, ..., f_r$, yielding (14).

Now pick an $f_l \in \{f_1, ..., f_r\}$. Then

$$\begin{aligned} f_l * h_m(w) &= \int_F f_l(z)\overline{h_m(z,-\overline{w})}y^k \ \frac{dxdy}{y^2} \\ &= \int_F f_l(z)\overline{\sum_{i,j=1}^{r} a_{ij} f_i(z) f_j(-\overline{w})}y^k \ \frac{dxdy}{y^2} \\ &= \sum_{i,j=1}^{r} \overline{a_{ij}} f_j(w) \int_F f_l(z)\overline{f_i(z)}y^k \ \frac{dxdy}{y^2}, \end{aligned}$$

using the fact that $f_j(w) = \overline{f_j(-\overline{w})}$ (we can prove this identity using the Fourier expansion of f_j). But the last integral is just the Petersson product of f_l and f_i, and therefore

$$\begin{aligned} f_l * h_m(w) &= \sum_{i,j=1}^{r} \overline{a_{ij}} f_j(w)\langle f_l, f_i\rangle \\ &= \sum_{j=1}^{r} \overline{a_{lj}} f_j(w)\langle f_l, f_l\rangle \end{aligned}$$

since the eigenforms are orthogonal. From Theorem 3.6, we also know that $f_l * h_m(w) = C_k m^{-k+1} a_m^{(l)} f_l(w)$, so it follows that

$$C_k m^{-k+1} a_m^{(l)} f_l(w) = \sum_{j=1}^{r} \overline{a_{lj}} f_j(w)\langle f_l, f_l\rangle.$$

Since the $f_1, f_2, ..., f_r$ are linearly independent, this result implies that $\overline{a_{lj}} = 0$ if $l \neq j$, and that otherwise

$$\overline{a_{ll}} = \frac{C_k m^{-k+1}}{\langle f_l, f_l\rangle} a_m^{(l)}. \tag{15}$$

In particular, $a_{ll} \in \mathbb{R}$ for each l. Returning to the expansion of h_m in (14), we find that

$$\begin{aligned} h_m(z,w) &= \sum_{l=1}^{r} a_{ll} f_l(z) f_l(w) \\ &= \sum_{l=1}^{r} \frac{C_k m^{-k+1}}{\langle f_l, f_l\rangle} a_m^{(l)} f_l(z) f_l(w), \end{aligned}$$

and therefore

$$C_k^{-1} m^{k-1} h_m(z,w) \;=\; \sum_{l=1}^{r} \frac{a_m^{(l)}}{\langle f_l, f_l \rangle} f_l(z) f_l(w)$$

as claimed. □

We now conclude the section with our preliminary trace formula for the Hecke operator $T(m)$.

Theorem 3.7. *For each positive integer m and even integer $k \geq 4$, the trace of the Hecke operator $T_k(m)$ acting on S_k is given by the formula*

$$Tr(T_k(m)) \;=\; C_k^{-1} m^{k-1} \int_F h_m(z, -\overline{z}) y^k \; \frac{dxdy}{y^2},$$

where F is a fundamental domain for $\Gamma = SL_2(\mathbb{Z})$ in $\mathbb{H}$.

Proof. Recall that the trace of an operator is the sum of its eigenvalues. In this case

$$Tr(T_k(m)) \;=\; \sum_{l=1}^{r} a_m^{(l)},$$

where the $a_m^{(l)}$ are defined as above. With this in mind, the proof is a straightforward application of Lemma 3.1. We have

$$\begin{aligned}
C_k^{-1} m^{k-1} \int_F h_m(z, -\overline{z}) y^k \; \frac{dxdy}{y^2} \;&=\; \int_F \sum_{l=1}^{r} \frac{a_m^{(l)}}{\langle f_l, f_l \rangle} f_l(z) f_l(-\overline{z}) \; \frac{dxdy}{y^2} \\
&=\; \sum_{l=1}^{r} \frac{a_m^{(l)}}{\langle f_l, f_l \rangle} \int_F f_l(z) \overline{f_l(z)} y^k \; \frac{dxdy}{y^2} \\
&=\; \sum_{l=1}^{r} \frac{a_m^{(l)}}{\langle f_l, f_l \rangle} \langle f_l, f_l \rangle \\
&=\; \sum_{l=1}^{r} a_m^{(l)} \\
&=\; Tr(T_k(m)),
\end{aligned}$$

where we've once again used the fact that $f_l(-\overline{z}) = \overline{f_l(z)}$. □

4 The Eichler-Selberg Trace Formula

In the previous section, we derived an expression for the trace of the Hecke operator $T(m)$ in terms of an integral over a fundamental domain of $SL_2(\mathbb{Z})$. Unfortunately, although this formula and the methods we used to find it may offer some insight, the result is not very useful. The function h_m is far from simple, and the reader may wonder if the integral in Theorem 3.7 is even computable.

The aim of this section is to analyze the integral

$$C_k^{-1} m^{k-1} \int_F h_m(z, -\overline{z}) y^k \, \frac{dxdy}{y^2} \tag{16}$$

and show that we can, in fact, convert it into a more manageable form. The key insight and motivation comes from a direct correspondence with the theory of binary quadratic forms, as we will see after an initial decomposition of (16). This connection will lead us to a useful and computable trace formula. Note that throughout this section, we let m be a positive integer and let $k \geq 4$ be an even integer.

4.1 Setting Up the Theorem

Plugging the value of $h_m(z, -\overline{z})$ into (16), we turn our attention to

$$\int_F \sum_{ad-bc=m} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k} \frac{dxdy}{y^2}, \tag{17}$$

where the expression in the denominator follows from the fact that k is even. This integral does not depend on the choice of domain F, so if $\gamma \in \Gamma = SL_2(\mathbb{Z})$ then

$$\begin{aligned} \int_F h_m(\gamma, -\overline{\gamma z}) y^k \, \frac{dxdy}{y^2} &= \int_{\gamma F} h_m(z, -\overline{z}) y^k \, \frac{dxdy}{y^2} \\ &= \int_F h_m(z, -\overline{z}) y^k \, \frac{dxdy}{y^2}. \end{aligned}$$

Therefore the sum inside (17) is invariant under the action of Γ, since otherwise the above equality would give us a contradiction. The plan is to use conjugacy classes to break this sum into chunks which are individually invariant under Γ.

So suppose that $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{M}^m$ and $\gamma = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right) \in \Gamma$, and write the conjugate $\gamma^{-1} M \gamma$ as $\left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right)$. A computation shows

$$c_2|z|^2 + d_2\overline{z} - a_2 z - b_2 \;=\; (c|\gamma z|^2 + d\overline{\gamma z} - a\gamma z - b)|c_1 z + d_1|^2,$$

which implies that

$$\frac{\Im(\gamma z)^k}{(c|\gamma z|^2 + d\overline{\gamma z} - a\gamma z - b)^k} \;=\; \frac{y^k}{(c_2|z|^2 + d_2\overline{z} - a_2 z - b_2)}$$

since $\Im(\gamma z) = |c_1 z + d_1|^{-2}\Im(z)$. This messy computation tells us that replacing z by γz in (17) is the same as replacing $M \in \mathbb{M}^m$ by the conjugate $\gamma^{-1}M\gamma$. In other words, letting $\gamma \in \Gamma$ act on z leads to a rearrangement of the order of summation of $h_m(z, -\overline{z})$.

We want to decompose h_m into pieces which are themselves invariant under Γ, and the conjugacy relation allows us to do this. Since the two matrices M and $\gamma^{-1}M\gamma$ have the same trace for all $\gamma \in \Gamma$, we can collect classes of $\mathbb{M}^m$ whose elements have trace t and write

$$\int_F h_m(z, -\overline{z})y^k \;\frac{dxdy}{y^2} \;=\; \sum_{t=-\infty}^{\infty} \int_F \sum_{\substack{ad-bc=m \\ a+d=t}} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k} \;\frac{dxdy}{y^2}. \quad (18)$$

Then the value of

$$\sum_{\substack{ad-bc=m \\ a+d=t}} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k}$$

is invariant under the action of Γ on z. The sum is over conjugacy classes whose elements have the same trace, and we've just seen that for any $\gamma \in \Gamma$, replacing z by γz is the same as replacing M by $\gamma^{-1}M\gamma$ in the sum.

With this in mind, if we set

$$I(m,t) \;=\; C_k^{-1} m^{k-1} \int_F \sum_{\substack{ad-bc=m \\ a+d=t}} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k} \;\frac{dxdy}{y^2},$$

then by (18) and Theorem 3.7 we have

$$Tr(T(m)) \;=\; \sum_{t=-\infty}^{\infty} I(m,t). \quad (19)$$

The reader may wonder why we didn't just begin with this decomposition, since on the surface it seems pretty simple. The point of all our work above was to show that the sum in each integral $I(m,t)$ is invariant under Γ, a fact which will be useful later on.

We now turn our attention to calculating $I(m,t)$ for an arbitrary $t \in \mathbb{Z}$. We use a correspondence between matrices in $\mathbb{M}^m$ with trace t and binary quadratic forms, made precise in the following proposition. This result will lead us to the main theorem of the paper, which provides a computable formula for the trace of $T(m)$.

Recall that a **binary quadratic form** (in integers) is a homogeneous polynomial of degree 2 in two variables, say

$$g(u,v) \;=\; au^2 + buv + cv^2$$

with $a, b, c \in \mathbb{Z}$. The **discriminant** of g is the value $|g| = b^2 - 4ac$.

Proposition 4.1. *Suppose $t \in \mathbb{Z}$. There is a one-to-one correspondence between matrices in $\mathbb{M}^m$ with trace t and binary quadratic forms with discriminant $t^2 - 4m$.*

Proof. First, suppose $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{M}^m$ with trace $t = a + d$, and construct with its entries a binary quadratic form $g(u,v) = cu^2 + (d-a)uv - bv^2$. Then

$$\begin{aligned} |g| &= (d-a)^2 + 4bc \\ &= a^2 + d^2 - 2ad + 4bc \\ &= t^2 - 4ad + 4bc \\ &= t^2 - 4m, \end{aligned}$$

as required.

Conversely, suppose we have a binary quadratic form $g(u,v)$ with integer coefficients and discriminant $t^2 - 4m$. We let $g(u,v) = \alpha^2 + \beta uv + \gamma v^2$, where $\beta = \pm t$ and $\alpha\gamma = m$. Then the matrix

$$M \;=\; \begin{pmatrix} \frac{t-\beta}{2} & -\gamma \\ \alpha & \frac{t+\beta}{2} \end{pmatrix}$$

has determinant m, trace t, and integer entries.

So we have an invertible mapping between matrices in $\mathbb{M}^m$ with trace t and binary quadratic forms with discriminant $t^2 - 4m$, yielding the desired one-to-one correspondence. □

We can now reinterpret our earlier decomposition of the trace formula. The idea is to think of the sum inside $I(m,t)$ in terms of matrix representations of binary quadratic forms with discriminant $t^2 - 4m$. More specifically, suppose we have an arbitrary form $g(u,v) = \alpha u^2 + \beta uv + \gamma v^2$ with $\alpha, \beta, \gamma \in \mathbb{Z}$. For $t = \pm\beta$ and $z = x + iy \in \mathbb{H}$, we set

$$\begin{aligned} R_g(z,t) \;&:=\; \frac{y^k}{\Big(\alpha(x^2+y^2) + \gamma + (\frac{t+\beta}{2})(x - iy) + (\frac{t-\beta}{2})(x+iy)\Big)^k} \\ &=\; \frac{y^k}{\big(\alpha(x^2+y^2) + \beta x + \gamma - ity\big)^k}. \end{aligned}$$

Here we've written the summand inside (19) in terms of the matrix representation of g. With this in mind, Proposition 4.1 allows us to write

$$I(m,t) \;=\; C_k^{-1} m^{k-1} \int\limits_F \sum_{|g|=t^2-4m} R_g(z,t)\; \frac{dxdy}{y^2}, \tag{20}$$

where the sum is taken over all binary quadratic forms with discriminant t^2-4m.

We want to analyze the sum inside (20), so the rest of our work will draw from some basic properties of binary quadratic forms (BQFs). In particular, recall that $\Gamma = SL_2(\mathbb{Z})$ acts on binary quadratic forms as follows: if $B = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \Gamma$ and $g(u,v)$ is a binary quadratic form, we let $Bg := g(pu+qv, ru+sv)$. Then Bg is a BQF with the same discriminant as g, and we say that a BQF f is **equivalent** to g if there is some $\vartheta \in \Gamma$ such that $f = \vartheta g$. As expected, the action of Γ gives us an equivalence relation on the set of binary quadratic forms.

This brief overview guides us to the following lemma:

Lemma 4.1. *If $B \in \Gamma$ and g is a binary quadratic form, then $R_g(Bz,t) = R_{Bg}(z,t)$.*

Proof. Direct computation. □

For the sake of notation, let $D = t^2 - 4m$. If a binary quadratic form g has discriminant D, we know that any equivalent form will also have discriminant D, and we use this fact to break up the sum inside (20). We write $|g| = D \pmod{\Gamma}$ to say that g is a representative of an equivalence class of binary quadratic forms with discriminant D. Additionally, we let Γ_g be the isotropy group (or stabilizer) of g, that is, the set of all elements of Γ that fix g. Then

$$\sum_{|g|=D} R_g(z,t) \;=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \sum_{B \in \Gamma/\Gamma_g} R_{Bg}(z,t),$$

where the inner sum is taken over left cosets of Γ_g in Γ. We add over these cosets to ensure that we don't count the same BQF more than once. The choice of each representative in Γ/Γ_g doesn't matter, since if $C, A \in \Gamma$ and $C \in A\Gamma_g$, then $C = A\gamma$ for some $\gamma \in \Gamma_g$, and hence $Cg = Ag$.

Lemma 4.1 now tells us that

$$\sum_{|g|=D} R_g(z,t) \;=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \sum_{B \in \Gamma/\Gamma_g} R_g(Bz,t),$$

so by (20)

$$C_k m^{-k+1} I(m,t) \;=\; \int\limits_F \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \sum_{B \in \Gamma/\Gamma_g} R_g(Bz,t)\; \frac{dxdy}{y^2}.$$

Let $h(D)$ denote the number of equivalence classes of binary quadratic forms with discriminant D. We call $h(D)$ the **class number** of such forms. If $D \neq 0$ then $h(D)$ is finite, so in that case

$$C_k m^{-k+1} I(m,t) \;=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \int_F \sum_{B \in \Gamma/\Gamma_g} R_g(Bz,t)\ \frac{dxdy}{y^2}.$$

Since we can bound $\sum R_g(Bz,t) = \sum R_{Bg}(z,t)$ by the function $h_m(z,-\overline{z})$, the sum of $R_g(Bz,t)$ over Γ/Γ_g clearly converges absolutely and uniformly. We can therefore exchange the order of integration and summation, giving us

$$\begin{aligned} C_k m^{-k+1} I(m,t) \;&=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \left(\sum_{B \in \Gamma/\Gamma_g} \int_{BF} R_g(z,t)\ \frac{dxdy}{y^2} \right) \\ &=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2}, \end{aligned} \tag{21}$$

where $F_g = \bigcup_{B \in \Gamma/\Gamma_g} BF$ for some choice of representatives. Note that F_g is a fundamental domain for Γ_g.

This is as far as we go before stating the theorem for the case $D \neq 0$. The case $D = 0$ is a slight variation. First note that we can take $\{g_r(u,v) = rv^2 \mid r \in \mathbb{Z}\}$ as a complete set of representatives of binary quadratic forms with discriminant 0. Additionally, if $r \neq 0$ then the isotropy group Γ_{g_r} of g_r is

$$\Gamma_\infty \;:=\; \left\{ \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \;\middle|\; n \in \mathbb{Z} \right\},$$

which we can easily verify: if $\gamma \in \Gamma$ and $\gamma g_r = g_r$, then the bottom row of γ must be $(0 \quad \pm 1)$ by definition, and since γ has determinant 1, this forces the top row to be $(\pm 1 \quad n)$ for some $n \in \mathbb{Z}$. On the other hand, if $r = 0$ then clearly $\Gamma_{g_r} = \Gamma$. It's worth noting that $\Gamma_\infty = \{\pm T^n \mid n \in \mathbb{Z}\}$, where $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is one of the two matrices that generates Γ.

Now if we let

$$R_0 \;:=\; \int_F R_{g_0}(z,t)\ \frac{dxdy}{y^2},$$

then proceeding as before we have

$$\int_F \sum_{|g|=0} R_g(z,t)\ \frac{dxdy}{y^2} \;=\; R_0 \;+\; \int_{F_\infty} \sum_{r \in \mathbb{Z}^*} R_{g_r}(z,t)\ \frac{dxdy}{y^2}, \tag{22}$$

where $F_\infty = \bigcup_{B \in \Gamma/\Gamma_\infty} BF$ is a fundamental domain for Γ_∞. Note that we cannot bring the inner sum out of the integral in this case: it is easy to check that

$$\int_{F_\infty} R_{g_r}(z,t)\frac{dxdy}{y^2} = 0$$

for all $r \in \mathbb{Z}^*$, but we show in Section 4.3, Case 2 that the integral of the sum in (22) is non-zero.

4.2 Statement of the Theorem

Before we can write down our main theorem, which gives the Eichler-Selberg Trace Formula for level-one Hecke operators, we need a few more definitions.

We define a function $H(n)$ on integers n by first setting

$$H(n) = 0 \text{ if } n < 0 \text{ and } H(0) = -\frac{1}{12}.$$

If $n > 0$, we let $H(n)$ be the number of equivalence classes (with respect to $\Gamma = SL_2(\mathbb{Z})$) of positive definite binary quadratic forms with discriminant $-n$, counting forms equivalent to a multiple of $u^2 + v^2$ or $u^2 + uv + v^2$ with multiplicity $\frac{1}{2}$ or $\frac{1}{3}$, respectively. Recall that a **positive definite** form $g(u,v)$ is greater than 0 for all pairs $(u,v) \neq (0,0)$. For such a form

$$g(u,v) \;=\; au^2 + buv + cv^2,$$

the discriminant D equals $b^2 - 4ac$, so clearly $D \equiv 0$ or $1 \pmod 4$. This implies that $-D \equiv 0$ or $3 \pmod 4$, and therefore $H(n) = 0$ if $n \equiv 1$ or $2 \pmod 4$. In the case $n > 0$, $H(n)$ is a fraction of the class number $h(-n)$ and is therefore always finite. If Γ_g is the stabilizer of a binary quadratic form g with discriminant D as before, then we also have

$$H(n) \;=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma) \\ g \text{ pos. def.}}} \frac{1}{|\Gamma_g|}.$$

We should note that the value of $H(n)$ is known for millions of inputs.

Next, for $t \in \mathbb{Z}$ and integers $m \geq 1$ and $j \geq 2$, we let $P_j(t,m)$ be the coefficient of x^{j-2} in the power series expansion of $(1 - tx + mx^2)^{-1}$. In other words, we have

$$(1 - tx + mx^2)^{-1} \;=\; \sum_{j=2}^{\infty} P_j(t,m)x^{j-2}.$$

The following lemma gives us a more concrete definition of $P_j(t,m)$.

Lemma 4.2. *Let ρ_1 and ρ_2 be the roots of the polynomial $x^2 - tx + m = 0$, and let $D = t^2 - 4m$. If $D \neq 0$ then*

$$P_j(t,m) \;=\; \frac{\rho_1^{j-1} - \rho_2^{j-1}}{\rho_1 - \rho_2}.$$

On the other hand, if $D = 0$ then

$$P_j(t,m) \;=\; \pm(j-1)m^{(j-1)/2},$$

where m is a square.

Proof. We let

$$\rho_1 = \frac{t + \sqrt{D}}{2} \quad \text{and} \quad \rho_2 = \frac{t - \sqrt{D}}{2}.$$

Note that we have the identities $\rho_1\rho_2 = m$ and $\rho_1 - \rho_2 = \sqrt{D}$.

First suppose $D \neq 0$, so that $\rho_1 \neq \rho_2$. Then we can write

$$\begin{aligned}
(1 - tx + mx^2)^{-1} \;&=\; \frac{\rho_1 - \rho_2}{\sqrt{D}(1 - \rho_1 x)(1 - \rho_2 x)} \\
&=\; \frac{2\rho_1(1 - \rho_2 x) - 2\rho_2(1 - \rho_1 x)}{2\sqrt{D}(1 - \rho_1 x)(1 - \rho_2 x)} \\
&=\; \frac{t + \sqrt{D}}{2\sqrt{D}}\left(\frac{1}{1 - \rho_1 x}\right) - \frac{t - \sqrt{D}}{2\sqrt{D}}\left(\frac{1}{1 - \rho_2 x}\right).
\end{aligned}$$

Using the geometric expansion, we get

$$\begin{aligned}
(1 - tx + mx^2)^{-1} \;&=\; \sum_{j=0}^{\infty} \rho_1^{j+1}\frac{x^j}{\sqrt{D}} \;-\; \sum_{j=0}^{\infty} \rho_2^{j+1}\frac{x^j}{\sqrt{D}} \\
&=\; \sum_{j=0}^{\infty} \frac{\rho_1^{j+1} - \rho_2^{j+1}}{\rho_1 - \rho_2} x^j,
\end{aligned}$$

and therefore

$$P_j(t,m) \;=\; \frac{\rho_1^{j-1} - \rho_2^{j-1}}{\rho_1 - \rho_2}.$$

This proves the first case. Now if $D = t^2 - 4m = 0$, then

$$\begin{aligned}
(1 - tx + mx^2)^{-1} \;&=\; \left(1 - tx + \frac{t^2}{4}x^2\right)^{-1} \\
&=\; \frac{1}{\left(1 - \frac{t}{2}x\right)^2} \\
&=\; \sum_{j=0}^{\infty}(j+1)\left(\frac{t}{2}\right)^{j+1} x^j.
\end{aligned}$$

Using the fact that $t = \pm 2\sqrt{m}$, with m a square, we conclude that

$$P_j(t,m) \;=\; \pm(j-1)m^{(j-1)/2}.$$

□

We can now state the main result of the paper.

Theorem 4.1 (Eichler-Selberg Trace Formula)**.** *Suppose m is a positive integer and $k \geq 4$ is an even integer. Then the trace of the Hecke operator $T(m)$ on the space S_k is given by*

$$Tr(T(m)) \;=\; -\frac{1}{2}\sum_{t=-\infty}^{\infty} P_k(t,m)H(4m-t^2) - \frac{1}{2}\sum_{dd'=m} \min\{d,d'\}^{k-1},$$

where d and d' are positive integers such that $dd' = m$.

We will use our results from Section 4.1 to tackle the proof of this formula. Recall that we can write

$$Tr(T(m)) \;=\; \sum_{t=-\infty}^{\infty} I(m,t),$$

and that we've analyzed the integral $I(m,t)$ in terms of binary quadratic forms. As before, let $D = t^2 - 4m$. We will prove that

$$\frac{1}{2}(I(m,t)+I(m,-t)) \;=\; \begin{cases} -\frac{1}{2}P_k(t,m)H(4m-t^2) & \text{if } D<0 \\ \frac{k-1}{48}m^{(k-1)/2} - \frac{1}{4}m^{(k-1)/2} & \text{if } D=0 \\ -\frac{1}{2}\left(\frac{|t|-u}{2}\right)^{k-1} & \text{if } D=u^2,\ u\in\mathbb{Z}^+ \\ 0 & \text{if } D>0 \\ & \text{and non-square,} \end{cases}$$

and then show that these equalities imply Theorem 4.1. In the wake of our results from the last section, each case is largely computational.

4.3 The Four Cases

A disclaimer: unless otherwise stated, the functions we use in the following cases all behave well enough to exchange the order of integration and summation, differentiation and integration, etc. This is mainly because we can always bound $\sum R_g(z,t)$ by $h_m(z,-\overline{z})$. We also use some algebraic and computational results from the theories of binary quadratic forms and quadratic fields. These facts are well-known, so we do not prove them.

Case 1: $D < 0$. By Section 4.1, we want to compute

$$\sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2},$$

where the sum runs over representatives g of classes of binary quadratic forms with discriminant D, and F_g is a fundamental domain for the stabilizer of each g. In this case the stabilizer Γ_g is a finite group, as discussed above in the definition of $H(n)$. Let

$$g(u,v) = \alpha u^2 + \beta uv + \gamma v^2$$

be a quadratic form with discriminant $|g| = D < 0$. Since F_g is a fundamental domain for Γ_g and the integral $I(m,t)$ does not depend on the choice of domain, we see that

$$\begin{aligned} \int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} &= \frac{1}{|\Gamma_g|} \int_{\mathbb{H}} R_g(z,t)\ \frac{dxdy}{y^2} \\ &= \frac{1}{|\Gamma_g|} \int_{\mathbb{H}} \frac{y^k}{(\alpha|z|^2 + \beta x + \gamma - ity)^k}\ \frac{dxdy}{y^2}. \end{aligned}$$

Using the substitution $z \mapsto \left(\frac{z}{\alpha} - \frac{\beta}{2\alpha}\right)$, we then have

$$\begin{aligned} \int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} &= \frac{1}{|\Gamma_g|} \int_{\mathbb{H}} \frac{y^k}{\alpha^k(\frac{1}{\alpha}|z|^2 + \frac{\beta^2}{4\alpha} - \frac{\beta^2}{2\alpha} + \gamma - it\frac{y}{\alpha})^k}\ \frac{dxdy}{y^2} \\ &= \frac{1}{|\Gamma_g|} \int_{\mathbb{H}} \frac{y^k}{(|z|^2 - ity - D/4)^k}\ \frac{dxdy}{y^2}, \end{aligned}$$

recalling that $\frac{dxdy}{y^2}$ is invariant under the action of $\left(\begin{smallmatrix} 2 & -\beta \\ 0 & 2\alpha \end{smallmatrix}\right) \in GL_2(\mathbb{Z})$. Now let

$$I := \frac{1}{|\Gamma_g|} \int_{\mathbb{H}} \frac{y^k}{(|z|^2 - ity - D/4)^k}\ \frac{dxdy}{y^2}.$$

The integral I depends only on t and D, so we can write

$$\begin{aligned} C_k m^{-k+1} I(m,t) &= \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} = \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \frac{1}{|\Gamma_g|}\ I \\ &= 2H(-D)I \qquad (23) \end{aligned}$$

using the second definition of H. The factor of 2 comes from the fact that in the definition of $H(n)$ we only counted positive definite forms, but here we also count their negatives.

In order to determine the value of $I(m,t)$, we will compute the integral I. We begin with the following lemma.

Lemma 4.3. *If $\Re(s) \neq 0$ and $l \geq 2$, then*

$$\int_{-\infty}^{\infty} (x^2+s)^{-l}\, dx = \frac{\pi}{(l-1)!}\frac{1}{2}\cdot\frac{3}{2}\cdots\left(l-\frac{3}{2}\right) s^{-l+1/2}$$

Proof. First consider the case $l = 2$. Then

$$\begin{aligned}\int_{-\infty}^{\infty} (x^2+s)^{-2}\, dx &= \frac{1}{2s^2}\int_{-\infty}^{\infty} \frac{2}{x^2/s+1} - \frac{2x^2/s}{(x^2/s+1)^2}\, dx \\ &= \frac{1}{2s^{3/2}}\left[\frac{x/\sqrt{s}}{(x^2/s)+1} + \tan^{-1}\left(\frac{x}{\sqrt{s}}\right)\right]_{-\infty}^{\infty} \\ &= \frac{\pi s^{-3/2}}{2}\end{aligned}$$

whenever $\Re(s) \neq 0$. Taking derivatives with respect to s, we have

$$\begin{aligned}\left(\frac{d}{ds}\right)^j \int_{-\infty}^{\infty} (x^2+s)^{-2}\, dx &= \left(\frac{d}{ds}\right)^j \frac{\pi s^{-3/2}}{2} \\ &= \frac{\pi}{2}\cdot\frac{-3}{2}\cdots\left(-\frac{1}{2}-j\right) s^{-(j+3/2)}.\end{aligned}$$

Since $\Re(s) \neq 0$, the function $(x^2+s)^{-2}$ is holomorphic and we can therefore bring the derivative $\left(\frac{d}{ds}\right)^j$ into the integral on the left. This proves the lemma. □

Applying the lemma, we see that

$$\begin{aligned}I &= \int_0^{\infty} y^{k-2} \int_{-\infty}^{\infty} (x^2+y^2-ity-D/4)^{-k}\, dxdy \\ &= \frac{\pi}{(k-1)!}\frac{1}{2}\cdot\frac{3}{2}\cdots(k-3/2)\int_0^{\infty}(y^2-ity-D/4)^{-k+1/2}\, y^{k-2}\, dy \\ &= \frac{\pi i^{k-2}}{2(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\int_0^{\infty}(y^2-ity-D/4)^{-3/2}\, dy \\ &= \frac{\pi i^{k-2}}{2(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\left(\frac{4}{t^2-D}\left[\frac{y-i\frac{t}{2}}{\sqrt{y^2-ity-D/4}}\right]_{y=0}^{\infty}\right)\end{aligned}$$

by completing the square and using the substitution $(y - i\frac{t}{2}) \mapsto \frac{1}{2}\sqrt{D+t^2}\tan(s)$. Therefore

$$\begin{aligned} I &= \frac{\pi i^{k-2}}{2(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\left(\frac{4}{t^2-D}+\frac{4}{t^2-D}\frac{it}{\sqrt{-D}}\right) \\ &= \frac{\pi i^{k-2}}{2(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\left(\frac{4}{\sqrt{|D|}}\frac{1}{\sqrt{|D|}-it}\right) \\ &= \frac{2\pi}{k-1}\frac{1}{\sqrt{|D|}}\frac{1}{(\sqrt{|D|}-it)^{k-1}}, \end{aligned}$$

noting that $-D = |D|$. Combining this result with (23), we find that

$$I(m,t) = C_k^{-1} m^{k-1} 2H(4m-t^2)\frac{2\pi}{k-1}\frac{1}{\sqrt{4m-t^2}}\frac{1}{(\sqrt{4m-t^2}-it)^{k-1}} \quad (24)$$

The important work is done; the rest of the proof is just a matter of simplifying this expression. We first write

$$\begin{aligned} &\frac{m^{k-1}}{\sqrt{4m-t^2}}\frac{1}{(\sqrt{4m-t^2}-it)^{k-1}}\cdot\frac{(\sqrt{4m-t^2}+it)^{k-1}}{(\sqrt{4m-t^2}+it)^{k-1}} \\ &\qquad = \frac{1}{\sqrt{4m-t^2}}\frac{(\sqrt{4m-t^2}+it)^{k-1}}{4^{k-1}} \\ &\qquad = \frac{(-1)^{k/2}(-i)^{k-1}}{i\sqrt{4m-t^2}}\frac{(\sqrt{4m-t^2}+it)^{k-1}}{4^{k-1}} \\ &\qquad = \frac{(-1)^{k/2}}{i\sqrt{4m-t^2}}\frac{(-i\sqrt{4m-t^2}+t)^{k-1}}{4^{k-1}}, \end{aligned} \quad (25)$$

using the fact that $i = (-1)^{k/2}(-i)^{k-1}$ since k is even. Now let $\rho = \frac{1}{2}(t + i\sqrt{4m-t^2})$. Putting (25) and (24) together and simplifying the constants, we get

$$\begin{aligned} I(m,t) &= H(4m-t^2)\cdot\frac{\frac{1}{2^{k-1}}(t-i\sqrt{4m-t^2})^{k-1}}{i\sqrt{4m-t^2}} \\ &= \frac{(\overline{\rho})^{k-1}}{\rho-\overline{\rho}}H(4m-t^2). \end{aligned}$$

If we think of ρ as a function of t, then a quick computation shows that $\rho(-t) = -\overline{\rho}(t)$. With this in mind, we can finish the case:

$$\begin{aligned}\frac{1}{2}(I(m,t)+I(m,-t)) &= -\frac{1}{2}\left(\frac{\rho^{k-1}-\overline{\rho}^{k-1}}{\rho-\overline{\rho}}H(4m-t^2)\right)\\ &= -\frac{1}{2}P_k(t,m)H(4m-t^2)\end{aligned}$$

as claimed.

Case 2: $D = 0$. In this case $t^2 = 4m$, and therefore $t = \pm\sqrt{m}$ with m a square. From Section 4.1, we know that when $D = 0$,

$$\begin{aligned}I(m,t) &= C_k^{-1}m^{k-1}\int_F \sum_{|g|=0} R_g(z,t)\ \frac{dxdy}{y^2}\\ &= C_k^{-1}m^{k-1}\left(\int_F R_{g_0}(z,t)\ \frac{dxdy}{y^2} + \int_{F_\infty}\sum_{r\in\mathbb{Z}^*} R_{g_r}(z,t)\ \frac{dxdy}{y^2}\right). \quad (26)\end{aligned}$$

As before, $g_r = rv^2$ with $r \in \mathbb{Z}$, and F_∞ is a fundamental domain for $\Gamma_\infty = \{\pm\left(\begin{smallmatrix}1 & n\\ 0 & 1\end{smallmatrix}\right) \mid n \in \mathbb{Z}\}$ in $\mathbb{H}$. In particular, we can take $F_\infty = \{z \in \mathbb{H} \mid 0 \le \Re(z) \le 1\}$.

We want to compute the two integrals in (26). The first is easy, since

$$R_{g_0}(z,t) = \frac{y^k}{(-ity)^k},$$

and therefore

$$\begin{aligned}\int_F R_{g_0}(z,t)\ \frac{dxdy}{y^2} &= \frac{(-1)^{k/2}}{t^k}\int_F \frac{dxdy}{y^2}\\ &= \frac{(-1)^{k/2}}{t^k}\frac{\pi}{6}.\end{aligned}$$

On to the second integral. We have

$$\begin{aligned}\int_{F_\infty}\sum_{r\in\mathbb{Z}^*} R_{g_r}(z,t)\ \frac{dxdy}{y^2} &= \int_0^\infty\int_0^1 \sum_{r\in\mathbb{Z}^*}\frac{y^k}{(r-ity)^k}\ \frac{dxdy}{y^2}\\ &= \int_0^\infty y^{k-2}\sum_{r\in\mathbb{Z}^*}(r-ity)^{-k}\ dy\\ &= \frac{i^{k-2}}{(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\int_0^\infty\sum_{r\in\mathbb{Z}^*}(r-ity)^{-2}\ dy. \quad (27)\end{aligned}$$

Now recall that for $\tau \in \mathbb{C}$, we can write

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\tau)^2} \;=\; \frac{\pi^2}{\sin^2(\pi\tau)}.$$

This identity is well-known, and its proof can be found in many analysis textbooks. Therefore

$$\begin{aligned}\sum_{r\in\mathbb{Z}} \frac{1}{(r-ity)^2} \;&=\; \frac{\pi^2}{\sin^2(\pi(-ity))} \\ &=\; -\frac{\pi^2}{\sinh^2(\pi ty)},\end{aligned}$$

since $\sin(-i\tau) = -i\sinh(\tau)$ for $\tau \in \mathbb{C}$. As a consequence we have

$$\sum_{r\in\mathbb{Z}^*} \frac{1}{(r-ity)^2} \;= \frac{1}{t^2y^2} - \frac{\pi^2}{\sinh^2(\pi ty)},$$

and therefore

$$\int_0^\infty \sum_{r\in\mathbb{Z}^*} (r-ity)^{-2}\, dy \;=\; \int_0^\infty \left(\frac{1}{t^2y^2} - \frac{\pi^2}{\sinh^2(\pi ty)}\right) dy. \tag{28}$$

The goal now is to compute the integral on the right side of (28). First, note that

$$\begin{aligned}-\frac{\pi^2}{\sinh^2(\pi ty)} \;&=\; -\pi^2 \mathrm{csch}^2(\pi ty) \\ &=\; \frac{d}{dy}\,\frac{\pi\coth(\pi ty)}{t},\end{aligned}$$

which implies that

$$\int_0^\infty \left(\frac{1}{t^2y^2} - \frac{\pi^2}{\sinh^2(\pi ty)}\right) \;=\; \lim_{\substack{R\to\infty \\ \epsilon\to 0}} \left[-\frac{1}{t^2y} + \frac{\pi\coth(\pi ty)}{t}\right]_{y=\epsilon}^{y=R} \tag{29}$$

where R and ϵ are positive real numbers. Using the identity

$$\coth(x) \;=\; \frac{e^{2x}+1}{e^{2x}-1},$$

we can apply L'Hopital's rule to show that $\coth(x) \to 1$ as $x \to \infty$. Therefore

$$\lim_{R\to\infty} -\frac{1}{t^2R} + \frac{\pi\coth(\pi tR)}{t} \;=\; \frac{\pi}{|t|}, \tag{30}$$

where the absolute value comes from the fact that $\coth(-z) = -\coth(z)$.

Next, using the Taylor series expansion of $\coth(x)$ around 0, we find that

$$\coth(x) = \frac{1}{x} + \frac{x}{3} + O(x^3)$$

as $x \to 0$. Therefore

$$\begin{aligned} -\frac{1}{t^2\epsilon} + \frac{\pi \coth(\pi t\epsilon)}{t} &= -\frac{1}{t^2\epsilon} + \frac{\pi}{t}\left(\frac{1}{\pi t\epsilon} + \frac{1}{3}(\pi t\epsilon) + O((\pi t\epsilon)^3)\right) \\ &= \frac{1}{3}(\pi^2 t\epsilon) + O(\epsilon^3), \end{aligned}$$

and as a consequence

$$\lim_{\epsilon \to 0} -\frac{1}{t^2\epsilon} + \frac{\pi \coth(\pi t\epsilon)}{t} = 0. \tag{31}$$

Putting (30) and (31) together, we find that

$$\lim_{\substack{R\to\infty \\ \epsilon\to 0}} \left[-\frac{1}{t^2 y} + \frac{\pi \coth(\pi t y)}{t}\right]_{y=\epsilon}^{y=R} = \frac{\pi}{|t|},$$

and hence with (28) and (29) we get

$$\int_0^\infty \sum_{r\in\mathbb{Z}^*} (r - ity)^{-2}\, dy = \frac{\pi}{|t|}.$$

Referring back to (27), we compute

$$\frac{i^{k-2}}{(k-1)!}\left(\frac{d}{dt}\right)^{k-2} \frac{\pi}{|t|} = \frac{(-1)^{(k-2)/2}\,\pi}{k-1}\,|t|^{-k+1}.$$

We've found the values of each integral in (26). Gathering our results, we get

$$\begin{aligned} I(m,t) &= C_k^{-1} m^{k-1}\left(\frac{(-1)^{k/2}}{t^k}\frac{\pi}{6} + \frac{(-1)^{(k-2)/2}\,\pi}{k-1}\,|t|^{-k+1}\right) \\ &= \frac{(-1)^{k/2}\pi}{2^{k-3}(k-1)} m^{k-1}\left(\frac{(-1)^{k/2}\,\pi}{4m^{k/2}\,6} + \frac{(-1)^{(k-2)/2}\,\pi}{(k-1)\,2^{k-1}m^{(k-1)/2}}\right) \\ &= \frac{(k-1)}{48} m^{(k-2)/2} - \frac{1}{4} m^{(k-1)/2}, \end{aligned}$$

using the fact that $t = \pm 2\sqrt{m}$ with m a square. The value of $I(m,t)$ doesn't depend on the sign of t, and therefore we're done.

Case 3: $D = u^2$ for some integer $u > 0$. If

$$g(u,v) = \alpha u^2 + \beta uv + \gamma v^2$$

is such a form with discriminant D, then it is easy to check that $|\Gamma_g| = 1$ (the Pell equation $x^2 + (uy)^2 = 4$ has only the trivial solution in integers). The set-up here is identical to the beginning of Case 1: since Γ_g is finite, we once again find that

$$\int_F \sum_{|g|=D} R_g(z,t)\ \frac{dxdy}{y^2} \ = \ HI,$$

where

$$H \ = \ \sum_{\substack{|g|=D \\ \bmod\ \Gamma}} \frac{1}{|\Gamma_g|} \quad \text{and} \quad I \ = \int_{\mathbb{H}} \frac{y^k}{(|z|^2 - ity - D/4)^k}\ \frac{dxdy}{y^2}$$

as above. We know from the theory of binary quadratic forms that there are u classes of BQFs with discriminant u^2, so it follows that $H = u$. Additionally, if we use the same methods as in the case $D < 0$, we get

$$I \ = \ \frac{\pi i^{k-2}}{2(k-1)!}\left(\frac{d}{dt}\right)^{k-2}\left(\frac{4}{t^2-D}\left[\frac{y - it/2}{\sqrt{y^2 - ity - D/4}}\right]_{y=0}^{\infty}\right).$$

We've just cheated a little bit, since in this case we have to worry about zeros of $(|z|^2 - ity - D/4)$ in $\mathbb{H}$ when we compute the integral I (e.g. the point $z = \frac{u}{2} + ity$ with $t > 0$). If we want to avoid any problems, we have to pay careful attention to what happens to our function around these zero points. It turns out that we can proceed as we did in Case 1, see [Wang and Pei] for the complete proof.

We also have to be careful when we compute

$$\frac{4}{t^2-D}\left[\frac{y - it/2}{\sqrt{y^2 - ity - D/4}}\right]_{y=0}^{\infty}, \tag{32}$$

since in this case we have to choose the branch of the square root with positive real part as we take the limit $y \to 0$. A computation similar to that in Case 1 shows that this expression equals

$$\frac{-4}{\sqrt{D}}\frac{1}{\sqrt{D} + |t|},$$

so that (32) only depends on $|t|$. Further, continuing just as we did in Case 1, we find that

$$HI \ = \ (-1)^{(k-2)/2}\frac{2\pi}{k-1}\frac{1}{(u+|t|)^{k-1}},$$

which implies that

$$I(m,t) \ = \ C_k^{-1} m^{k-1} HI \ = \ -\frac{1}{2}\left(\frac{|t| - u}{2}\right)^{k-1}.$$

This completes the case, since here $I(m,t) = I(m,-t)$.

Case 4: $D > 0$ and D is non-square. In this case the stabilizer Γ_g is infinitely cyclic, as we will soon show directly. The guiding intuition for this part of the theorem is that if Γ_g is an infinite set, then

$$H \;=\; \sum_{\substack{|g|=D \\ (\mathrm{mod}\ \Gamma)}} \frac{1}{|\Gamma_g|} \;= \sum \frac{1}{\infty} \;= 0,$$

so that $HI = 0$, where

$$I \;=\; \int_{\mathbb{H}} R_g(z,t)\ \frac{dxdy}{y^2}$$

as before.

To prove the case more precisely, we need to write Γ_g in a concrete form. So suppose we have a binary quadratic form $g(u,v) = \alpha u^2 + \beta uv + \gamma v^2$ with discriminant $\beta^2 - 4\alpha\gamma = D$, and let $w > w'$ be the roots of the equation $\alpha x^2 + \beta x + \gamma = 0$. For the sake of the argument we assume, without loss of generality, that

$$w \;=\; \frac{-\beta + \sqrt{D}}{2\alpha} \;\text{ and }\; w' \;=\; \frac{-\beta - \sqrt{D}}{2\alpha},$$

or in other words, that $\alpha > 0$. Now consider the matrix

$$J \;:=\; \frac{1}{\sqrt{w - w'}} \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix} \in SL_2(\mathbb{R}).$$

We will first show that if $T \in \Gamma_g$ and $T \neq \pm I$, then

$$T \;=\; J \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^{-1} \end{pmatrix} J^{-1}$$

for a real $\epsilon \neq \pm 1$. We will then show that this ϵ is in fact a unit in the ring of integers of $\mathbb{Q}(\sqrt{D})$, and hence a power of the number field's fundamental unit. As a consequence, Γ_g is cyclic and we can write down all of its elements.

Before continuing, we need the following lemma.

Lemma 4.4. *$Jg = -\sqrt{D}xy$. In other words,*

$$J^T \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix} J \;=\; -\frac{\sqrt{D}}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proof. We have

$$J^T \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix} J \;=\; \frac{1}{w - w'} \begin{pmatrix} w\alpha + \frac{\beta}{2} & \frac{w\beta}{2} + \gamma \\ w'\alpha + \frac{\beta}{2} & \frac{w'\beta}{2} + \gamma \end{pmatrix} \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}$$

$$= \frac{1}{w - w'} \begin{pmatrix} \frac{\sqrt{D}}{2} & \frac{\beta\sqrt{D}}{4\alpha} \\ -\frac{\sqrt{D}}{2} & \frac{-\beta\sqrt{D}}{4\alpha} \end{pmatrix} \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}$$

using the above values of w and w'. We also know that $w - w' = \frac{\sqrt{D}}{\alpha}$, which implies that

$$J^T \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix} J \;=\; \begin{pmatrix} \frac{\alpha}{2} & \frac{\beta-\sqrt{D}}{4} \\ -\frac{\alpha}{2} & \frac{-\beta-\sqrt{D}}{4} \end{pmatrix} \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}.$$

One more computation shows that the right hand side is

$$-\frac{\sqrt{D}}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

as claimed. □

Suppose now that $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_g$ and that $T \neq \pm I$. Then

$$T^T \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix} T \;=\; \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix}.$$

If we let $S = J^{-1}TJ \neq \pm I \in SL_2(\mathbb{R})$, then

$$\begin{aligned} S^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} S \;&=\; J^T T^T (J^{-1})^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} J^{-1}TJ \\ &=\; J^T T^T (J^{-1})^T J^T \frac{-2}{\sqrt{D}} \begin{pmatrix} \alpha & \beta/2 \\ \beta/2 & \gamma \end{pmatrix} JJ^{-1}TJ \end{aligned}$$

using the above lemma. Simplifying this expression and applying the lemma once again, we get

$$S^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} S \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This is only possible if

$$S \;=\; \begin{pmatrix} 0 & a \\ a^{-1} & 0 \end{pmatrix} \quad \text{or} \quad S \;=\; \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

for some $a \in \mathbb{R}$, as the reader can check. But the $(1,1)$-entry of S is nonzero, so it follows that there is an $\epsilon \in \mathbb{R}$ such that

$$S \;=\; \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^{-1} \end{pmatrix}.$$

We also know that $\epsilon \neq \pm 1$, since this would imply that $T = \pm I$. It then follows that

$$T \;=\; J \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^{-1} \end{pmatrix} J^{-1}, \tag{33}$$

and hence that $Tz = \epsilon^2 z$ for $z \in \mathbb{H}$. Writing $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ as before and evaluating T in terms of (33), we find (after some basic calculations) that

$$\frac{aw+b}{cw+d} = w \quad \text{and} \quad \frac{aw'+b}{cw'+d} = w'.$$

This result implies that $-c(w)^2+(a-d)w+b = 0$ and $-c(w')^2+(a-d)w'+b = 0$, and therefore that w and w' are roots of the polynomial

$$p(x) = cx^2 + (d-a)x - b.$$

Writing

$$\frac{p(x)}{c} \;=\; (x-w)(x-w') \;=\; x^2 - x(w+w') + ww',$$

we find that $\frac{d-a}{c} = \frac{\beta}{\alpha}$ and $-\frac{b}{c} = \frac{\gamma}{\alpha}$ (since $w + w' = -\frac{\beta}{\alpha}$ and $ww' = \frac{\gamma}{\alpha}$). We cannot determine the value of c, so instead we write $c = r\alpha$ for some $r \in \mathbb{Q}$ such that $r\alpha \in \mathbb{Z}$. Then we have

$$c = r\alpha, \quad d - a = r\beta, \quad \text{and} \quad -b = r\gamma. \tag{34}$$

Now let $t = a + d$ be the trace of T, so that $\epsilon + \epsilon^{-1} = t$ by (33). We have

$$a = \frac{t - r\beta}{2}, \quad b = -r\gamma, \quad c = r\alpha, \quad d = \frac{t + r\beta}{2}, \tag{35}$$

and

$$t^2 - Dr^2 = (a+d)^2 - r^2(\beta^2 - 4\alpha\gamma) = 4,$$

which the reader can easily check. Since $\epsilon + \epsilon^{-1} = t$ and $\epsilon\epsilon^{-1} = 1$, we also know that ϵ and ϵ^{-1} are the roots of the polynomial $q(x) = x^2 - tx + 1$. Without loss of generality, it follows that

$$\epsilon^{\pm 1} \;=\; \frac{t \pm \sqrt{t^2-4}}{2} \;=\; \frac{t \pm r\sqrt{D}}{2}, \tag{36}$$

using the fact that $t^2 - Dr^2 = 4$.

Set $r = s/q$ with q and s integers, $\gcd(s,q) = 1$, and $q \geq 1$. Then from (34) we get $qc = s\alpha$, so that $q|s\alpha$. But $\gcd(s,q) = 1$, so this implies that $q|\alpha$. Likewise $q|\beta$ and $q|\gamma$, and as a consequence $q|\gcd(\alpha,\beta,\gamma)$. Additionally, since $D = \beta^2 - 4\alpha\gamma$, it is clear that $\gcd(\alpha,\beta,\gamma)^2|D$.

So let

$$D = \Delta \gcd(\alpha,\beta,\gamma)^2$$

with $\Delta \in \mathbb{N}$ a non-square. Since $q|\gcd(\alpha,\beta,\gamma)$, there is a $p' \in \mathbb{Z}$ such that $\gcd(\alpha,\beta,\gamma) = p'q$. If we set $p = sp'$, we get

$$r = \frac{p}{\gcd(\alpha,\beta,\gamma)}.$$

Therefore from (36),

$$\epsilon^{\pm 1} = \frac{t \pm \frac{p}{\gcd(\alpha,\beta,\gamma)}\sqrt{\Delta \gcd(\alpha,\beta,\gamma)^2}}{2}$$

$$= \frac{t \pm p\sqrt{\Delta}}{2}. \tag{37}$$

Moreover, since $t^2 - Dr^2 = 4$ we also have

$$t^2 - \Delta \gcd(\alpha,\beta,\gamma)^2 \frac{p^2}{\gcd(\alpha,\beta,\gamma)^2} = 4,$$

implying that

$$t^2 - \Delta p^2 = 4. \tag{38}$$

From quadratic field theory, we know that there is a bijection between integral solutions to the Pell equation in (38) and units in the ring of integers of the number field $\mathbb{Q}(\sqrt{\Delta})$. Additionally, we know that if δ is such a unit, then

$$\delta = \pm\epsilon_0^n, \quad 0 \neq n \in \mathbb{Z},$$

where

$$\epsilon_0 = \frac{t_0 + p_0\sqrt{\Delta}}{2} > 1$$

is the fundamental unit of the field (here (t_0, p_0) the smallest solution to (38)). But ϵ is a unit in $\mathbb{Q}(\sqrt{\Delta})$ by (37), and therefore

$$\epsilon = \pm\epsilon_0^n, \quad n \in \mathbb{Z}^*.$$

Using (33) we then find that

$$T = \pm J \begin{pmatrix} \epsilon_0^n & 0 \\ 0 & \epsilon_0^{-n} \end{pmatrix} J^{-1} = \pm \left(J \begin{pmatrix} \epsilon_0 & 0 \\ 0 & \epsilon_0^{-1} \end{pmatrix} J^{-1} \right)^n$$

for some $n \in \mathbb{Z}^*$, implying that Γ_g is the infinitely cyclic group generated by

$$J \begin{pmatrix} \epsilon_0 & 0 \\ 0 & \epsilon_0^{-1} \end{pmatrix} J^{-1}.$$

We now have enough material to compute $I(m,t)$. We first write

$$\int_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} = \int_{F_g} R_{Jg}(J^{-1}z,t)\ \frac{dxdy}{y^2} = \int_{J^{-1}F_g} R_{Jg}(z,t)\ \frac{dxdy}{y^2}. \tag{39}$$

Since Γ_g is generated by $J \left(\begin{smallmatrix} \epsilon_0 & 0 \\ 0 & \epsilon_0^{-1} \end{smallmatrix} \right) J^{-1}$, we can chose the fundamental domain F_g such that $J^{-1}F_g$ is the annulus $\Lambda = \{z = x + iy \mid y > 0,\ 1 < |z| \leq \epsilon_0^2\}$. (If $Jz \in F_g$, then $J \left(\begin{smallmatrix} \epsilon_0^n & 0 \\ 0 & \epsilon_0^{-n} \end{smallmatrix} \right) z = \epsilon^{2n} Jz$ is not in F_g).

Therefore, using (39) and Lemma 4.1, we see that

$$\begin{aligned}\int\limits_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} &= \int\limits_{\Lambda} R_{Jg}(z,t)\ \frac{dxdy}{y^2} \\ &= \int\limits_{\Lambda} R_{-\sqrt{D}xy}(z,t)\ \frac{dxdy}{y^2} \\ &= \int\limits_{\substack{y>0 \\ 1<|z|\leq\epsilon_0^2}} \frac{y^k}{(\sqrt{D}x+ity)^k}\ \frac{dxdy}{y^2}.\end{aligned}$$

Changing to polar coordinates, with $x=\rho\cos\theta$ and $y=\rho\sin\theta$, we then have

$$\begin{aligned}\int\limits_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} &= \int\limits_0^{\pi}\int\limits_1^{\epsilon_0^2} (\rho\sqrt{D}\cos\theta+\rho it\sin\theta)^{-k}(\rho\sin\theta)^{k-2}\rho\ d\rho d\theta \\ &= \int\limits_0^{\pi}\int\limits_1^{\epsilon_0^2} (\sqrt{D}\cos\theta+it\sin\theta)^{-k}(\sin\theta)^{k-2}\rho^{-1}\ d\rho d\theta \\ &= 2\log(\epsilon_0)\int\limits_0^{\pi}(\sqrt{D}\cos\theta+it\sin\theta)^{-k}(\sin\theta)^{k-2}\ d\theta.\end{aligned}$$

With this result in mind, we let

$$\begin{aligned}\Omega &:= \int\limits_{F_g} R_g(z,t)\ \frac{dxdy}{y^2} + \int\limits_{F_g} R_g(z,-t)\ \frac{dxdy}{y^2} \\ &= 2\log(\epsilon_0)\int\limits_{-\pi}^{\pi}(\sqrt{D}\cos\theta+it\sin\theta)^{-k}(\sin\theta)^{k-2}\ d\theta,\end{aligned}$$

noting that the $-t$ in the integral of $R_g(z,-t)$ corresponds to a change of variable $\theta\mapsto-\theta$. Now since we also have

$$\Omega = 2\log(\epsilon_0)\int\limits_{-\pi}^{\pi}\left(\sqrt{D}\frac{e^{i\theta}+e^{-i\theta}}{2}-t\frac{e^{i\theta}-e^{-i\theta}}{2}\right)^{-k}\left(\frac{e^{i\theta}-e^{-i\theta}}{2i}\right)^{k-2}\ d\theta,$$

we let $\zeta=e^{i\theta}$ and write Ω in terms of an integral on the unit disk:

$$\Omega = 8(-1)^{k/2}i\log(\epsilon_0)\oint\limits_{|\zeta|=1}\frac{(\zeta^2-1)^{k-2}\zeta}{((\sqrt{D}-t)\zeta^2+(\sqrt{D}+t))^k}\ d\zeta.$$

The function

$$f(z) \;=\; \frac{(z^2-1)^{k-2}z}{((\sqrt{D}-t)z^2+(\sqrt{D}+t))^k}$$

is holomorphic when $|z| \leq 1$, so $\Omega = 0$ by Cauchy's theorem. Therefore $I(m,t)+I(m,-t) = \sum \Omega = 0$, completing the final case.

4.4 Finishing the Proof

All we have to do now is tie these four cases together and show that they imply the Eichler-Selberg Trace Formula. We have two last cases to consider, when the integer m is a square and when it is not. Assuming, first, that m is a square, we turn our attention to the sum

$$\sum_{t=-\infty}^{\infty} I(m,t),$$

which gives the trace of the Hecke operator $T(m)$ as we showed in 4.1. Note that

$$\sum_{t=-\infty}^{\infty} \frac{1}{2}(I(m,t)+I(m,-t)) \;=\; \frac{1}{2}\left(\sum_{t=-\infty}^{\infty} I(m,t) + \sum_{t=-\infty}^{\infty} I(m,-t)\right)$$

$$= \; Tr(T(m)).$$

We can therefore think of $Tr(T(m))$ as a sum

$$\sum_{\substack{t \\ t^2-4m<0}} \Lambda_{m,t} \;+\; \sum_{\substack{t \\ t^2-4m=0}} \Lambda_{m,t} \;+\; \sum_{\substack{t \\ t^2-4m=u^2 \\ u>0 \,\in\, \mathbb{Z}}} \Lambda_{m,t} \;+\; \sum_{\substack{t \\ t^2-4m>0 \\ \text{non-square}}} \Lambda_{m,t},$$

with $\Lambda_{m,t} = \frac{1}{2}(I(m,t)+I(m,-t))$. By Case 4, the last term of this sum is zero, so we only need to consider the first three pieces. Note that we can freely rearrange the sum of $\Lambda_{m,t}$ since $I(m,t)$ is an integral of a piece of $h_m(z,w)$ and $h_m \in H_k^2(\Gamma)$ as a function of z.

Now by Case 1,

$$\sum_{\substack{t \\ t^2-4m<0}} \Lambda_{m,t} \;=\; -\frac{1}{2}\sum_{\substack{t=-\infty \\ t^2\neq 4m}}^{\infty} P_k(t,m)H(4m-t^2), \tag{40}$$

which is actually a finite sum since $H(4m-t^2) = 0$ for $t > 2\sqrt{m}$. If m is a square, there are two values of t for which $t^2 = 4m$, and therefore

$$\sum_{\substack{t \\ t^2-4m=0}} \Lambda_{m,t} \;=\; 2\left(\frac{k-1}{48}m^{(k-2)/2} - \frac{1}{4}m^{(k-1)/2}\right)$$

$$= \; \frac{k-1}{24}m^{(k-2)/2} - \frac{1}{2}m^{(k-1)/2} \tag{41}$$

using our result from Case 2. From Lemma 4.2, we know that

$$\frac{k-1}{24}m^{(k-2)/2} = -\frac{1}{2}\left(-\frac{1}{12}\right)P_k(2\sqrt{m}, m).$$

Since $H(0) = -\frac{1}{12}$, the first term in (41) is $-\frac{1}{2}P_k(2\sqrt{m}, m)H(0)$. Putting (40) and (41) together, we then have

$$\sum_{\substack{t \\ t^2-4m<0}} \Lambda_{m,t} \; + \sum_{\substack{t \\ t^2-4m=0}} \Lambda_{m,t} \; = \; -\frac{1}{2}\sum_{t=-\infty}^{\infty} P_k(t,m)H(4m-t^2) - \frac{1}{2}m^{(k-1)/2}.$$

To finish the proof of the theorem, observe that if $t^2 - 4m = u^2$ for some $u > 0 \in \mathbb{Z}$, then

$$\begin{aligned}\frac{|t|+u}{2} \cdot \frac{|t|-u}{2} \; &= \; \frac{t^2-u^2}{4} \\ &= \; \frac{t^2-(t^2-4m)}{4} \; = \; m,\end{aligned}$$

noting that if $t^2 - 4m = u^2$, then $|t| > u$. Since t and u must be either both even or both odd ($u = \pm\sqrt{t^2-4m} \in \mathbb{Z}$), this tells us that

$$\frac{|t|+u}{2} \quad \text{and} \quad \frac{|t|-u}{2}$$

are both positive integers and hence divisors of m. In particular, there can only be finitely many such u. Since $u > 0$, we also see that

$$\min\left(\frac{|t|+u}{2}\,,\,\frac{|t|-u}{2}\right) \; = \; \frac{|t|-u}{2}.$$

In this light, Case 3 tells us that

$$\sum_{\substack{t \\ t^2-4m=u^2 \\ u>0 \,\in\, \mathbb{Z}}} \Lambda_{m,t} \; = \; -\frac{1}{2}\sum_{\substack{dd'=m \\ d\neq d'}} \min(d,d')^{k-1}. \tag{42}$$

If $dd' = m$ and $d = d'$, then $d = d' = 2\sqrt{m}$. Therefore, combining (40), (41), and (42), we get

$$Tr(T(m)) \; = \; -\frac{1}{2}\sum_{t=-\infty}^{\infty} P_k(t,m)H(4m-t^2) - \frac{1}{2}\sum_{dd'=m} \min\{d,d'\}^{k-1}$$

as claimed. This completes the proof of the Eichler-Selberg Trace Formula in the case where m is a square. But the case where m is not a square only requires a slight adjustment: in this case it is not possible to have $t^2 - 4m = 0$, since this implies that m is a square. Therefore (40) and (42) are all we need to prove the theorem.

5 Conclusion

The most important point to make about this version of the Eichler-Selberg Trace Formula is that it's computable: we know the value of the $H(n)$ function for millions of inputs, Lemma 4.2 gives us a method to find the values of $P_k(t, m)$, and a computer can determine the value of

$$\sum_{dd'=m} \min(d, d')^{k-1}.$$

The E-S Formula also implies some relationships among the class numbers of binary quadratic forms. For example, when $k = 4$ the only cusp form is 0, so we have

$$\sum_{dd'=m} \min(d, d')^{k-1} \;=\; -\sum_{t=-\infty}^{\infty} P_k(t, m) H(4m - t^2)$$

for positive integers m.

Finally, once we have the trace of a Hecke operator $T_k(m)$, there are some methods we can use to actually compute the individual eigenvalues. These algorithms rely on the interesting algebraic properties of the Hecke operators that we briefly mentioned in Section 2. For one example, see the end of Section 6.8 in [Miyake].

References

[1] Flath, D. *Introduction to Number Theory*. Wiley-Interscience, 1989.

[2] Lang, S., *Introduction to Modular Forms*, Springer, 1976.

[3] Miyake, T. *Modular Forms*. Springer, 1989.

[4] Pei, D. and Wang, X., *Modular Forms with Integral and Half-Integral Weights*.Springer, Science Press Beijing, 2012.

[5] Serre, J-P. *A Course in Arithmetic*. Springer-Verlag, 1973.

[6] Zagier, D., Bruinier, J.H., van der Geer, G., and Harder, G. *The 1-2-3 of Modular Forms*. Springer, 2008.