Class Number Relations

1. Introduction

Let $H(n)$ denote the (weighted) number of classes of positive binary quadratic forms of discriminant $-n$:

$$H(n) = \sum_{f^2|n} \hat{h}(-n/f^2),$$

where

$$\hat{h}(\Delta) = \begin{cases} 2|Cl(\mathcal{O}_\Delta)|/|\mathcal{O}_\Delta^\times|, & \text{if } \Delta \equiv 0,\ 1 \pmod 4, \\ 0 & \text{otherwise,} \end{cases}$$

and put $H(0) = -\frac{1}{12}$. In addition, for $D \geq 1$ let

$$H_D(n) = \sum_{x \in S_D(n)} H\left(\frac{4n - x^2}{D}\right),$$

where $S_D(n) = \{x \in \mathbb{Z} : x^2 \leq 4n,\ x^2 \equiv 4n \pmod D\}$. Furthermore, for $D = N^2$ let

$$I_{N^2}(n) = \frac{1}{N} \sum_{\substack{d_1 d_2 = n \\ d_1 \equiv d_2\ (N)}} \min(d_1, d_2).$$

Theorem A (Kronecker, 1857; Gierster, 1879)

(1) $$H_1(n) + I_1(n) = 2\sigma(n), \quad \text{for all } n \geq 1.$$

Remarks. 1) Gierster (1880), Hurwitz (1884/5) study $H_{p^2}(n)$ (and other related functions) for $p = 7, 11, 13$.

2) In Dickson's *History of the Theory of Numbers*, there is a whole chapter (cf. vol. III, chapter VII) which is devoted to such class number relations.

Theorem B (Hurwitz, 1885) If $N = p$ is prime, then there is a cusp form $f(z) = \sum a_n(f)q_N^n \in S_2(\Gamma(N))$ (where $q_N = e^{2\pi iz/N}$) such that

(2) $$\frac{p(p^2-1)}{2}\left[H_{p^2}(n) + I_{p^2}(n)\right] = 2\sigma(n) + a_n(f),$$

whenever $\left(\frac{n}{p}\right) = 1$.

Theorem C (Hirzebruch-Zagier, 1976) If D is the discriminant of a real quadratic number field K, then

$$\varphi_D(z) := \sum_{n\geq 0}[H_D(n) + I_D(n)]q^n \in M_2(D, \chi_D),$$

where $$I_D(n) = \frac{1}{\sqrt{D}} \sum_{\substack{\lambda \in \mathcal{O}_K \\ \lambda >> 0 \\ \lambda\lambda' = n}} \min(\lambda, \lambda').$$

2. A Generalization of Hurwitz's Theorem

Aims. 1) To generalize Theorem B to arbitrary N.

2) To refine the theorem in such a way so as to give precise information about the error term $f \in S_2(\Gamma(N))$.

Notation. If $(k, N) = 1$, then let

$$I_{N^2,k}(n) = \frac{1}{N} \sum_{\substack{d_1 d_2 = n \\ d_1 \equiv d_2 \equiv \pm k\,(N)}} \min(d_1, d_2).$$

and

$$H_{N^2,k}(n) = \sum_{x \in S_{N^2,k}(n)} H\left(\frac{4n - x^2}{N^2}\right),$$

where

$$S_{N^2,k}(n) = \{x \in S_{N^2}(n) : x \equiv \pm 2k + \tfrac{4n-x^2}{N}\,(2N)\}.$$

Remark. Recall that

$$S_{N^2}(n) = \{x \in \mathbb{Z} : x^2 \le 4n,\ x^2 \equiv 4n\ (N^2)\}.$$

It is easy to see that $S_{N^2}(n) = \dot{\bigcup}_{\pm k} S_{N^2,k}(n)$. Thus

$$H_{N^2}(n) = \sum_{\pm k} H_{N^2,k}(n),\ I_{N^2}(n) = \sum_{\pm k} I_{N^2,k}(n).$$

Notation. Put

$$t_{N^2,k}(n) = \begin{cases} \tilde{H}_{N^2,k}(n) - \frac{2}{m}\sigma(n) & \text{if } n \equiv k^2 \ (N) \\ 0 & \text{otherwise,} \end{cases}$$

where $m = |SL_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}|$ and where

$$\tilde{H}_{N^2,k}(n) := H_{N^2,k}(n) + I_{N^2,k}(n).$$

Note that by definition

$$\tilde{H}_{N^2,k}(n) = H_{N^2,k}(n) = I_{N^2,k} = 0, \ \text{if } n \not\equiv k^2 \ (N).$$

Theorem 1 The function

$$\varphi_{N,k}(z) := \sum t_{N,k}(n) q_N^n$$

is a cusp form of weight 2 on $\Gamma(N)$.

Corollary 1. Theorem B holds for arbitrary $N \geq 3$: there is a cusp form $f(z) = \sum a_n(f) q_N^n \in S_2(\Gamma(N))$ such that

(3) $$m\left[H_{N^2}(n) + I_{N^2}(n)\right] = s\sigma(n) + a_n(f),$$

whenever n is an invertible square mod N, i.e. whenever $n \equiv k^2 \operatorname{mod}, (N)$, for some k with $(k, N) = 1$. Here $s = [(\mathbb{Z}/N\mathbb{Z})^\times : ((\mathbb{Z}/N\mathbb{Z})^\times)^2]$.

Corollary 2. If $N \leq 5$ and $(k, N) = 1$, then

$$H_{N^2,k}(n) + I_{N^2,k}(n) = \frac{2}{m}\sigma(n), \quad \forall n \equiv k^2 \ (N).$$

Proof. $\dim S_2(\Gamma(N)) = 0$ for $N \leq 5$.

Corollary 3. If $N = 6$ and $(k, N) = 1$, then

$$H_{N^2,k}(n) + I_{N^2,k}(n) = \frac{1}{36}\left(\sigma(n) - a_n\right), \forall n \equiv 1(6),$$

where $\sum a_n q_6^n = q_6 \prod (1 - q_6^{6n})^4 \in S_2(\Gamma(6))$.

Proof. $\dim S_2(\Gamma(6)) = 1$.

Example: $N = 6$:

n	$H'(n)$	$I'(n)$	$\sigma(n)$	a_n	$\tilde{H}'(n)$	$\sigma(n) - a_n$
1	-6	6	1	1	0	0
7	0	12	8	-4	12	12
13	0	12	14	2	12	12
19	0	12	20	8	12	12
25	-6	42	31	-5	36	36
31	24	12	32	-4	36	36
37	36	12	38	-10	48	48
43	24	12	44	8	36	36
49	-6	54	57	9	48	48

where $H'(n) := 36H_{36}(n), I'(n) := 36I_{36}(n)$, etc.

3. A General Formula (via A-L Theory)

Atkin-Lehner Theory for $V = V_N = S_2(\Gamma(N))$:

$$V = \bigoplus_{f\in\mathcal{N}(V)} V_f,$$

where $\mathcal{N}(V)$ = set of all normalized newforms (of all levels) on V

$$V_f = \oplus_{M|N_f}\mathbb{C}f(Mz), \text{ where } N_f = \text{level}(f).$$

Note. The Atkin-Lehner Theory on V is inherited from that of $S_2(\Gamma_1(N^2))$ via the map $V \hookrightarrow S_2(\Gamma_1(N^2))$ given by $f(z) \mapsto f(Nz)$.

In particular, we have that $N_f|N^2$.

Notation. 1) For $f, g \in \mathcal{N}(V)$ write

$$f \approx g \Leftrightarrow \exists\chi \bmod N\colon a_n(f) = \chi(n)a_n(g), \forall (n, N) = 1.$$

and put $V(f) = \sum_{g\approx f} V_g$. Thus

$$V = \bigoplus_{f\in\mathcal{N}(V)/\approx} V(f). \tag{4}$$

2) For $f = \sum a_n q_N^n \in V$ and $(k, N) = 1$ let

$$f^{(k)} = \sum_{n\equiv k(N)} a_n(f) q_N^n.$$

Theorem 2. If $(N,k)=1$, then

$$(5)\ \varphi_{N,k} = -\frac{2}{m} \sum_{f\in\mathcal{N}(V_N)/\approx} \dim V(f)\chi_f(k) f^{(k^2)},$$

where χ_f denotes the Nebentypus character of $f \in \mathcal{N}(V_N)$.

Corollary. If $N \geq 5$, $(N,k)=1$ and $n \equiv k^2(N)$, then

$$|\tilde{H}_{N^2,k}(n) - \frac{2}{m}\sigma(n)| \leq \frac{N-5}{6N} d(n)\sqrt{n}.$$

Proof. By Hasse/Weil/Eichler/Igusa we have

$$|a_n(f)| \leq d(n)\sqrt{n}, \quad \text{if } f \in \mathcal{N}(V),$$

and so by (4) and (5) we obtain

$$\begin{aligned} |t_{N,k}(n)| &\leq \frac{2}{m}(\dim V) d(n)\sqrt{n} \\ &= \left(\frac{N-6}{6N} + \frac{2}{m}\right) d(n)\sqrt{n}. \end{aligned}$$

Remarks. 1) Note that this bound is essentially independent of N, for $\frac{N-5}{6N} \leq \frac{1}{6}$.

2) If N is prime then $\#(\mathcal{N}(V_N)/\approx)$ is easily calculated:

N	7	11	13	17	19	23	29
$\dim V_N$	3	26	50	133	196	375	806
$\#\mathcal{N}(V_N)$	3	25	48	128	189	363	784
$\#(\mathcal{N}(V_N)/\approx)$	1	3	4	8	11	18	28

Thus, the above theorem explains (and generalizes) the fact that in the cases $N = 7$ resp. $N = 11$ Gierster and Hurwitz require 1 resp. 3 new "number theoretic functions" in their expressions of the error term in their class number formulae.

4. Method of Proof (Overview)

Basic Idea (Gierster/Hurwitz, 1880): Compute the number of coincidences of certain modular correspondences (introduced by F. Klein) in two ways:

1) by counting intersection points: $\rightarrow$ L.S.

2) by computing integrals (traces): $\rightarrow$ R.S.

Remark. The fact that these two methods yield the same answer is non-trivial:

For $N = 1$: Correspondence Principle (Chasles, 1855).

For $N > 5$: the Chasles/Brill/Noether Correspondence Principle is not applicable!

In fact, as a consequence of his studies of these class number formulae, Hurwitz discovered the following:

Theorem (Hurwitz,1886): Let X be a Riemann surface, α a correspondence on X with graph $\Gamma_\alpha \subset X \times X$ and Rosati conjugate α^*. Then the number of coincidences of α is given by the formula

$$(\Delta.\Gamma_\alpha) = \deg(\alpha) + \deg(\alpha^*) - \mathrm{trace}(\alpha + \alpha^* | \Omega(X)),$$

where $\Omega(X)$ denotes the space of holomorphic differentials on X.

Remark. Later generalizations of Hurwitz's formula:

1) The Lefschetz fixed point formula;

2) The Eichler/Selberg trace formula.

Application: Take

$X = X(N) = \Gamma(N)\backslash\mathfrak{H}^*$ modular curve of level N
$\alpha = \hat{T}_{n,k} = T_n\sigma_k$, a modular correspondence on X, where $n \equiv k^2 \pmod{N}$ and $(k,n) = 1$.

Note: 1) The action of T_n on $\Omega(X(N)) \simeq S_2(\Gamma(N))$ is given by the (usual) Hecke operator T_n.

2) The automorphism $\sigma_k \in \mathrm{Aut}(X)$ is defined by the matrix

$$\sigma_k = \begin{pmatrix} k^{-1} & 0 \\ 0 & k \end{pmatrix} \in G_N := \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}.$$

Then we have:

$$(\hat{T}_{n,k})^* = \hat{T}_{n,k}, \tag{6}$$

$$\deg(\hat{T}_{n,k}) = \sigma(n), \tag{7}$$

and so we still need to compute:

1) the number of coincidences $(\Delta.\hat{T}_{n,k})$,

2) the trace $\mathrm{tr}(\hat{T}_{n,k}|\Omega(X(N)))$.

5. Method of Proof (Local Computations)

Observation. If $n \neq \square$, then we can write

$$(8) \qquad (\Delta.\hat{T}_{n,k}) = (\Delta.\hat{T}_{n,k})_{fin} + (\Delta.\hat{T}_{n,k})_{inf},$$

by considering local intersection numbers:

$$(\Delta.\hat{T}_{n,k})_{fin} = \sum_{x \in X(N)_{fin}} (\Delta.\hat{T}_{n,k})_{(x,x)}$$

$$(\Delta.\hat{T}_{n,k})_{fin} = \sum_{x \in X(N)_{\infty}} (\Delta.\hat{T}_{n,k})_{(x,x)}.$$

In fact, such a decomposition (8) is possible for all n.

Theorem 3. For any $N \geq 3$ and $n \equiv k^2(N)$ with $(N,k) = 1$ we have

$$(9) \qquad (\Delta.\hat{T}_{n,k})_{fin} = mH_{N^2,k}(n)$$

$$(10) \qquad (\Delta.\hat{T}_{n,k})_{inf} = mI_{N^2,k}(n)$$

Remark. Eichler (1967) states similar formulae, but his are incorrect if $8|N$ because he uses the (larger) set

$$S'_{N^2,k}(n) := \{x \in S_{N^2}(n) : x \equiv \pm 2k \,(\mathrm{mod}\, N)\}.$$

in place of $S_{N^2,k}(n)$. Note that $S'_{N^2,k}(n) = S_{N^2,k}(n)$ if N is odd, but that in general $S'_{N^2,k}(n) \neq S_{N^2,k}(n)$.

Proof Sketch. a) Use the modular interpretation:

$$\Gamma(N)\backslash\mathfrak{H} = X(N)_{fin} \ni x \qquad \leftrightarrow \qquad \text{isom.cl.}(E, \lambda)$$

where $E/\mathbb{C}$ is an elliptic curve,

$\lambda : E[N] \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2$ is a level-N-structure.

Then: 1) $(\Delta.\hat{T}_{n,k})_{(x,x)} \neq 0$

$\Rightarrow$ E has CM (complex multiplication),

$\Rightarrow$ $\text{End}(E) \simeq \mathcal{O}_d \subset \mathbb{Q}(\sqrt{d})$, for some $d < 0$.

2) If $q_E := \deg_E$ denotes the positive definite binary quadratic form of discriminant d defined by the degree on $\text{End}(E) \simeq \mathcal{O}_d \simeq \mathbb{Z}^2$, then

$$(\Delta.\hat{T}_{n,k})_{(x,x)} = \#\{(a,b) \in \mathbb{Z}^2 : q_E(a,b) = n, \\ (a,b) \equiv \pm(0,k) \pmod{N}\}$$

3) From this, together with the fact that

$$\#\{E : \text{End}(E) = \mathcal{O}_d\} = \hat{h}(d),$$

one obtains the first equation of Theorem 3 by using a suitable identity.

Proof Sketch (cont'd). b) Use the fact that at the point (∞, ∞), the curve $\hat{T}_{n,k}$ has a local equation of the form

$$\prod_{\substack{ad=n \\ a\equiv d\equiv \pm k(N)}} f_{a,d}(x,y) = 0,$$

where $f_{a,d} = y^d - x^a$. Since

$$(\Delta.(f_{a,d} = 0))_{(\infty,\infty)} = \min(a, d),$$

we obtain

$$(\Delta.\hat{T}_{n,k})_{(\infty,\infty)} = NI_{N^2,k}(n).$$

From this the assertion (10) follows because

$$(\Delta.\hat{T}_{n,k})_{(x,x)} = (\Delta.\hat{T}_{n,k})_{(\infty,\infty)}, \quad \forall x \in X(N)_\infty,$$

(and because $\#X(N)_\infty = \frac{m}{N}$.)

6. Method of Proof (Trace Computation)

Observation. The following groups and algebras act on the space $\Omega(X) \simeq V = S_2(\Gamma(N))$:

$G = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ – as automorphism group
$\mathbb{T} = \mathbb{C}[\ldots T_n \ldots]$ – the Hecke algebra
$\mathbb{M} = \langle \mathbb{T}, G \rangle$ – the algebra of modular correspondences

Theorem 4. a) Each $V(f)$ is an irreducible $\mathbb{M}$-module.

b) The centre of $\mathbb{M}$ is $Z(\mathbb{M}) = \mathbb{C}[\ldots \hat{T}_{n,k} \ldots]$.

Corollary. If $f \in \mathcal{N}(V)$, then:

$$\mathrm{trace}(\hat{T}_{n,k}|V(f)) = (\dim V(f))\chi_f(k)a_n(f).$$

Proof. By Theorem 4 (and Schur's lemma) we have

$$(\hat{T}_{n,k})_{|V(f)} = c \cdot id_{V(f)}, \text{ for some } c \in \mathbb{C},$$

and so $\mathrm{trace}(\hat{T}_{n,k}|V(f)) = c \dim V(f)$. On the other hand,

$$cf = f|\hat{T}_{n,k} = f|T_n\sigma_k = a_n(f)f|\sigma_k = a_n(f)\chi_f(k)f,$$

so $c = a_n(f)\chi_f(k)$, and hence the assertion follows.

Proof of Theorem 2. If $n \equiv k^2 \pmod N$, then by Hurwitz's trace formula and Theorem 3 we have

$$\begin{aligned}
a_n(\varphi_{N^2,k}) &= -\frac{2}{m}\mathrm{trace}(\hat{T}_{n,k}|V) \\
&\stackrel{(4)}{=} -\frac{2}{m}\sum_{f\in\mathcal{N}(V)/\approx} \mathrm{trace}(\hat{T}_{n,k}|V(f)) \\
&\stackrel{Cor}{=} -\frac{2}{m}\sum_{f\in\mathcal{N}(V)/\approx} \dim(V(f))\chi_f(k)a_n(f) \\
&= a_n\left(-\frac{2}{m}\sum_{f\in\mathcal{N}(V)/\approx}\dim V(f)\chi_f(k)f^{(k^2)}\right).
\end{aligned}$$

On the other hand, if $n \not\equiv k^2 \bmod N)$, then

$$a_n(\varphi_{N^2,k}) = 0 = a_n\left(-\frac{2}{m}\sum_{f\in\mathcal{N}(V)/\approx}\dim V(f)\chi_f(k)f^{(k^2)}\right)$$

by definition. Thus

$$\varphi_{N^2,k} = -\frac{2}{m}\sum_{f\in\mathcal{N}(V)/\approx}\dim V(f)\chi_f(k)f^{(k^2)} \in V,$$

as claimed.