On singular moduli.

Gross, B.H.

in: Journal für die reine und angewandte Mathematik - 355 | Periodical
30 page(s) (191 - 220)

On singular moduli

Dedicated to J-P. Serre

By *Benedict H. Gross* at Providence and *Don B. Zagier* at Bonn

---

0. The values of the modular function $j(\tau)$ at imaginary quadratic arguments τ in the upper half plane are known as singular moduli. They are all algebraic integers. In this paper we will study the prime factorizations of the differences of two singular moduli. These differences turn out to be highly divisible numbers. For instance, we will determine the set of primes dividing the absolute norm of $j(\tau)-1728=j(\tau)-j(i)$ (and the multiplicities with which they occur); they turn out to be contained among the prime divisors of the positive integers of the form $d-x^2$, where d is the discriminant of τ, and hence smaller than or equal to d, e.g.:

$$j\left(\frac{1+i\sqrt{163}}{2}\right)-1728=-2^6\,3^6\,7^2\,11^2\,19^2\,127^2\,163.$$

1. Let $j(\tau)$ denote the elliptic modular function on the upper half plane $\mathfrak{H}$. This is a holomorphic function which is invariant under the action of the modular group $\Gamma=PSL_2(\mathbb{Z})$ and has a Fourier expansion

$$(1.1)\qquad j(\tau)=\frac{1}{q}+744+196\,884\,q+21\,493\,760\,q^2+\cdots=\frac{1}{q}+\sum_{n\geq 0}c_n q^n$$

with $q=e^{2\pi i\tau}$ and $c_n\in\mathbb{Z}$ for all n.

As a function on $\mathfrak{H}$, $j(\tau)$ enjoys the following remarkable property:

Whenever τ lies in an imaginary quadratic extension $K=\mathbb{Q}(\tau)$ of the rational number field, $j(\tau)$ is an algebraic integer in an abelian extension of K. If $a\tau^2+b\tau+c=0$ where a, b and c are integers with g.c.d. $(a, b, c)=1$, and we define

$$d=\mathrm{disc}(\tau)=b^2-4ac,$$

then $j(\tau)$ is an integer of degree $h=h(d)$ over $\mathbb{Q}$. Here $h(d)$ is the class number of primitive binary quadratic forms of discriminant d, or equivalently, the order of the class group of the order $\mathbb{Z}\left[\frac{b+\sqrt{d}}{2}\right]$. The conjugates of $j(\tau)$ are the h values $j(\tau')$, where τ' ranges over all roots of primitive quadratic polynomials of discriminant d. Finally, the field $H=K(j(\tau))$ is abelian over K and "dihedral" over $\mathbb{Q}$; it is the ring class field of conductor f over K, where $d=d_K f^2$ ([3]).

Now suppose d_1 and d_2 are two fundamental discriminants which are relatively prime, and define $D=d_1 d_2$. Let w_1 and w_2 be the number of roots of unity in the

quadratic orders of discriminants d_1 and d_2 respectively, and let h_1 and h_2 denote their class numbers. Consider the product

$$J(d_1, d_2) = \left(\prod_{\substack{[\tau_1],[\tau_2] \\ \operatorname{disc} \tau_i = d_i}} (j(\tau_1) - j(\tau_2)) \right)^{\frac{4}{w_1 w_2}} \tag{1.2}$$

where $[\tau_i]$ denotes an equivalence class modulo Γ. Note that this is the absolute norm of the algebraic integer $j(\tau_1) - j(\tau_2)$ of degree $h_1 h_2$, provided that $d_1, d_2 < -4$. In those cases, $J(d_1, d_2)$ is an integer which depends only on d_1 and d_2; in general, $J(d_1, d_2)^2$ is an integer and our first result is a specific formula for it.

For primes l with $\left(\frac{d_1 d_2}{l}\right) \neq -1$ we define

$$\varepsilon(l) = \begin{cases} \left(\frac{d_1}{l}\right) & \text{if } (l, d_1) = 1, \\ \left(\frac{d_2}{l}\right) & \text{if } (l, d_2) = 1. \end{cases}$$

If $n = \prod_i l_i^{a_i}$ with $\left(\frac{D}{l_i}\right) \neq -1$ for all i, we define $\varepsilon(n) = \prod_i \varepsilon(l_i)^{a_i}$.

Theorem 1.3. $J(d_1, d_2)^2 = \pm \prod_{\substack{x, n, n' \in Z \\ n, n' > 0 \\ x^2 + 4nn' = D}} n^{\varepsilon(n')}$.

Note that $\varepsilon(n')$ is well-defined, for if l divides $\frac{D - x^2}{4}$ then $\left(\frac{D}{l}\right) \neq -1$. Also note that since $\varepsilon\left(\frac{D - x^2}{4}\right) = -1$ we could replace $\varepsilon(n')$ in the formula by $-\varepsilon(n)$. In fact, one can replace $\varepsilon(n')$ by $\operatorname{sign}(m)\left(\frac{d_1}{m}\right) = \operatorname{sign}(m)\left(\frac{d_2}{m}\right)$, where m is any integer prime to D which is represented by the binary quadratic form $[n', x, -n]$.

If neither d_1 nor d_2 equals -4, the sign in 1.3 is always $+$.

We may rewrite 1.3 in the form

$$J(d_1, d_2)^2 = \pm \prod_{\substack{x^2 < D \\ x^2 \equiv D \pmod 4}} F\left(\frac{D - x^2}{4}\right), \tag{1.4}$$

where

$$F(m) = \prod_{\substack{nn' = m \\ n, n' > 0}} n^{\varepsilon(n')}. \tag{1.5}$$

An interesting fact about $F(m)$ is that it is either 1 or the power of a single prime l. The latter case occurs if l is the unique prime dividing m to odd exponent with $\varepsilon(l) = -1$. More precisely, if

$$m = l^{2a+1} l_1^{2a_1} \cdots l_s^{2a_s} q_1^{b_1} \cdots q_r^{b_r}$$

with $\varepsilon(l) = \varepsilon(l_i) = -1$, $\varepsilon(q_i) = +1$, then $F(m) = l^{(a+1)(b_1+1)\cdots(b_r+1)}$. In particular, we have

Corollary 1.6. *If l is a rational prime dividing $J(d_1, d_2)^2$ then $\left(\frac{d_1}{l}\right) \neq 1$, $\left(\frac{d_2}{l}\right) \neq 1$, and l divides a positive integer of the form $\frac{D - x^2}{4}$. In particular, $l \leqq \frac{D}{4}$; if $D \equiv 1 \pmod 8$ then $l < \frac{D}{8}$ and if $d_1 \equiv d_2 \equiv 5 \pmod 8$ then $l < \frac{D}{16}$.*

As an illustration of the theorem, we take $d_1 = -67$, $d_2 = -163$ (the last two discriminants with class number 1). Then

$$J(-67, -163) = j\left(\frac{1+i\sqrt{67}}{2}\right) - j\left(\frac{1+i\sqrt{163}}{2}\right) = -2^{15}3^3 5^3 11^3 + 2^{18}3^3 5^3 23^3 29^3$$
$$= 2^{15} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 139 \cdot 331,$$

while $F\left(\frac{D-x^2}{4}\right)$ for $|x|$ odd and less than $\sqrt{d_1 d_2} = 104.5\ldots$ is given by the following table:

$\lvert x\rvert$	$\frac{D-x^2}{4}$	$F\left(\frac{D-x^2}{4}\right)$	$\lvert x\rvert$	$\frac{D-x^2}{4}$	$F\left(\frac{D-x^2}{4}\right)$	$\lvert x\rvert$	$\frac{D-x^2}{4}$	$F\left(\frac{D-x^2}{4}\right)$
1	$2\cdot3\cdot5\cdot7\cdot13$	1	35	$2^3\cdot3\cdot101$	1	69	$2^2\cdot5\cdot7\cdot11$	1
3	$2^3\cdot11\cdot31$	1	37	$2^2\cdot3\cdot199$	3^2	71	$2\cdot3\cdot5\cdot7^2$	1
5	$2^2\cdot3\cdot227$	3^2	39	$2\cdot5^2\cdot47$	2^2	73	$2\cdot3\cdot233$	1
7	$2\cdot3^2\cdot151$	2^2	41	$2\cdot3\cdot5\cdot7\cdot11$	1	75	$2^2\cdot331$	331
9	$2\cdot5\cdot271$	1	43	$2^2\cdot3^4\cdot7$	7	77	$2^5\cdot3\cdot13$	1
11	$2^2\cdot3^3\cdot5^2$	3^2	45	$2^4\cdot139$	139	79	$2\cdot3^2\cdot5\cdot13$	1
13	$2^7\cdot3\cdot7$	1	47	$2\cdot3^2\cdot11$	2	81	$2\cdot5\cdot109$	1
15	$2\cdot7\cdot191$	1	49	$2\cdot3\cdot5\cdot71$	1	83	$2^4\cdot3^2\cdot7$	7
17	$2\cdot3\cdot443$	1	51	$2^2\cdot5\cdot13$	1	85	$2^2\cdot3\cdot7\cdot11$	1
19	$2^4\cdot3\cdot5\cdot11$	1	53	$2^2\cdot3\cdot13^2$	3	87	$2\cdot419$	2^2
21	$2^2\cdot5\cdot131$	5^2	55	$2\cdot3\cdot329$	1	89	$2\cdot3\cdot5^3$	1
23	$2\cdot3\cdot433$	1	57	$2\cdot7\cdot137$	1	91	$2^2\cdot3\cdot5\cdot11$	1
25	$2\cdot3^2\cdot11\cdot13$	1	59	$2^2\cdot3\cdot5\cdot31$	1	93	$2^3\cdot71$	2^4
27	$2^2\cdot7^2\cdot13$	13	61	$2^3\cdot3^2\cdot5^2$	2^2	95	$2\cdot3\cdot79$	1
29	$2^3\cdot3^2\cdot5\cdot7$	1	63	$2\cdot11\cdot79$	1	97	$2\cdot3^3\cdot7$	1
31	$2\cdot3\cdot5\cdot83$	1	65	$2\cdot3^3\cdot31$	1	99	$2^3\cdot5\cdot7$	1
33	$2\cdot1229$	2^2	67	$2^3\cdot3\cdot67$	1	101	$2^2\cdot3^2\cdot5$	5
						103	$2\cdot3\cdot13$	1

(The large frequency of x with $F\left(\frac{D-x^2}{4}\right)=1$ is due to the fact that $F(n)=1$ whenever $l^{\text{odd}} \| n$ for at least two primes l with $\left(\frac{-163}{l}\right)=-1$, and $\left(\frac{-163}{l}\right)=-1$ for all $l<40$.)

The cases of Theorem 1.3 when $d_1 = -3$ or -4 give formulas for the norms of $j(\tau)^{\frac{2}{3}}$ and $(j(\tau)-1728)$. We have tabulated the results for all known fundamental discriminants with $h=1$ or $h=3$ (and for one discriminant each with $h=5$, 7) in the table on the next page (Table 1). They agree with the computations of Berwick [2], who in 1928 computed j for all known discriminants with $h \leqq 3$ and gave the ful factorization of j and $j-1728$ in the appropriate quadratic or cubic field.

Besides tabulating the prime factors, Berwick made several conjectures on congruences satisfied by $j(\tau)$ and $j(\tau)-1728$. We will prove all of these divisiblities; for example:

Table 1. *Factorizations of* $N_{Q(j)/Q}(j) = \pm a^3$, $N_{Q(j)/Q}(j-1728) = \pm b^2 d$, disc (min. poly. of j) $= I^2 d^{\frac{h-1}{2}}$

$\lvert d\rvert$	h	a	b	I
3	1	0	$2^3 3$	1
4	1	$2^2 3$	0	1
7	1	$3 \cdot 5$	3^3	1
8	1	$2^2 5$	$2^2 7$	1
11	1	2^5	$2^3 7$	1
19	1	$2^5 3$	$2^3 3^3$	1
43	1	$2^6 3 \cdot 5$	$2^3 3^4 7$	1
67	1	$2^5 3 \cdot 5 \cdot 11$	$2^3 3^3 7 \cdot 31$	1
163	1	$2^6 3 \cdot 5 \cdot 23 \cdot 29$	$2^3 3^3 7 \cdot 11 \cdot 19 \cdot 127$	1
23	3	$5^3 11 \cdot 17$	$7^3 11^2 19$	$5^9 7^6 11^2 17 \cdot 19$
31	3	$3^3 11 \cdot 17 \cdot 23$	$3^{10} 11^2$	$3^{19} 11^2 13^3 17 \cdot 23 \cdot 29$
59	3	$2^{16} 11$	$2^9 11^2 23 \cdot 43$	$2^{46} 11^2 13^3 23 \cdot 31 \cdot 47$
83	3	$2^{16} 5^3$	$2^9 19 \cdot 47 \cdot 67 \cdot 79$	$2^{46} 5^9 13^3 19 \cdot 47 \cdot 71$
107	3	$2^{15} 5^3 17$	$2^9 7^3 43 \cdot 71 \cdot 103$	$2^{48} 5^9 7^6 17 \cdot 31 \cdot 59 \cdot 71$
139	3	$2^{16} 3^3 23$	$2^9 3^{11} 103$	$2^{47} 3^{20} 17^3 19^3 23 \cdot 59$
211	3	$2^{17} 3^3 17 \cdot 29$	$2^9 3^9 7^3 23 \cdot 67$	$2^{46} 3^{21} 7^6 17 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 131 \cdot 167 \cdot 191$
283	3	$2^{15} 3^3 5^3 53$	$2^9 3^{10} 19^2 31 \cdot 139$	$2^{48} 3^{19} 5^9 17^3 19^2 31 \cdot 47 \cdot 107 \cdot 167 \cdot 191 \cdot 239$
307	3	$2^{17} 3^3 5^3 47$	$2^9 3^{11} 23 \cdot 163 \cdot 271$	$2^{49} 3^{19} 5^9 13^3 23 \cdot 29 \cdot 31 \cdot 47 \cdot 59 \cdot 61 \cdot 131 \cdot 239 \cdot 263$
331	3	$2^{15} 3^3 11 \cdot 23 \cdot 29 \cdot 59$	$2^9 3^{11} 7^3 11^2 59^2$	$2^{48} 3^{20} 7^6 11^2 13^3 23 \cdot 29 \cdot 41 \cdot 47 \cdot 59^2 151 \cdot 251 \cdot 263 \cdot 311$
379	3	$2^{17} 3^3 11 \cdot 17 \cdot 53 \cdot 71$	$2^9 3^9 7^4 11^2 31 \cdot 47^2$	$2^{46} 3^{21} 7^6 11^2 13^3 17 \cdot 29 \cdot 31 \cdot 43^2 47^2 59 \cdot 71 \cdot 89 \cdot 199 \cdot 359$
499	3	$2^{16} 3^3 17 \cdot 23 \cdot 41 \cdot 71 \cdot 83$	$2^9 3^{11} 7^3 71^2 463$	$2^{46} 3^{19} 7^6 11^6 13^3 17 \cdot 19^3 23 \cdot 41 \cdot 59 \cdot 71^2 113 \cdot 179 \cdot 311 \cdot 383 \cdot 419 \cdot 479$
547	3	$2^{15} 3^3 5^3 17 \cdot 23 \cdot 101$	$2^9 3^{11} 7^3 31^2 59 \cdot 223$	$2^{49} 3^{21} 5^{10} 7^6 17 \cdot 23 \cdot 31^2 41 \cdot 59 \cdot 71 \cdot 79 \cdot 83 \cdot 89 \cdot 101 \cdot 151 \cdot 359 \cdot 431 \cdot 503$
643	3	$2^{15} 3^3 5^3 11 \cdot 17^3 113$	$2^9 3^{11} 11^2 43 \cdot 67 \cdot 71 \cdot 499 \cdot 607$	$2^{50} 3^{19} 5^9 11^2 13^3 17^9 19^3 41^3 47 \cdot 59 \cdot 71 \cdot 109 \cdot 179 \cdot 239 \cdot 311 \cdot 431$
883	3	$2^{15} 3^3 5^3 11^2 41 \cdot 89 \cdot 113$	$2^9 3^{11} 7^3 11 \cdot 23 \cdot 43^2 307 \cdot 739$	$2^{48} 3^{20} 5^{10} 7^6 11^3 19^3 23 \cdot 41 \cdot 43^2 47^3 59^3 61 \cdot 89 \cdot 101 \cdot 151 \cdot 167 \cdot 173 \cdot 271 \cdot 359 \cdot 419 \cdot 599$
907	3	$2^{19} 3^3 5^3 131 \cdot 137 \cdot 167$	$2^9 3^9 7^3 47 \cdot 67^2 79 \cdot 331$	$2^{55} 3^{21} 5^9 7^6 11^6 17^3 29^3 31 \cdot 47 \cdot 59 \cdot 67^2 79 \cdot 83 \cdot 101 \cdot 149 \cdot 179 \cdot 223 \cdot 251 \cdot 439 \cdot 479 \cdot 743$
47	5	$5^5 11^2 23 \cdot 29$	$11^3 19^2 23^2 31 \cdot 43$	$5^{30} 11^9 13^{10} 19^5 23^3 29^2 31^2 41^2 43$
71	7	$11^3 17^2 23 \cdot 41 \cdot 47 \cdot 53$	$7^7 11^4 23^2 31^2 67$	$7^{42} 11^{21} 13^{21} 17^{13} 23^8 31^6 41^3 47^3 53^2 59^3 61^2 67$

$$(1.7)\qquad \begin{cases} \text{If } |d| \equiv 3 \ (8) & \text{then } j(\tau) \equiv 0 \bmod 2^{15}. \\ \text{If } |d| \equiv 1 \ (3) & \text{then } j(\tau) \equiv 1728 \bmod 3^6. \end{cases}$$

We shall also present some refinements and generalizations of theorem 1. 3 in the course of the paper. When $d_1 = d_2$ it is natural to replace the absolute norm of $j(\tau_1) - j(\tau_2)$, which is equal to zero, by the discriminant of the integral polynomial satisfied by $j(\tau)$. We will give a formula for the discriminant, which, in turn, determines the index I of the order $\mathbb{Z}[j]$ in its integral closure. This index, also given in Table 1, grows rapidly with d; for example when $d = -71$, so $h = 7$, our formula gives the value $7^{42}11^{21}13^{21}17^{13}23^{8}31^{6}41^{3}47^{3}53^{2}59^{3}61^{2}67$, a number which was found by McKay and Ford by a computer calculation ([15], p. 349). In general, we show all prime factors l of the discriminant are less than or equal to $|d|$. Finally, we shall prove some results on $\varphi_m(x, y)$, where φ_m is the m-th modular polynomial, at singular moduli. For $m = 1$ we have $\varphi_1(x, y) = x - y$, so this generalizes our previous results. In particular, we will show that any prime dividing $\varphi_m(j(\tau_1), j(\tau_2))$ has residue characteristic $l \leqq \frac{m^2 d_1 d_2}{4} = \frac{m^2 D}{4}$; more precisely, l must divide $\frac{m^2 D - x^2}{4}$ for some x with $|x| < m\sqrt{D}$ (see Table 2 for $m = 2$).

Table 2. *Factorization of* $\varphi_2(j_1, j_2)$ *for* $0 > d_1, d_2 > -20$, $h_1 = h_2 = 1$

d_1	d_2	$\varphi_2(j_1, j_2)$
−3	−3	$-2^{12}3^{9}5^{9}$
−4	−3	$-2^{12}3^{9}11^{6}$
−4	−4	0
−7	−3	$-3^{9}5^{9}17^{3}$
−7	−4	$-3^{20}7^{3}19^{2}$
−7	−7	0
−8	−3	$-2^{12}5^{9}23^{3}$
−8	−4	$2^{13}7^{6}23^{2}31^{2}$
−8	−7	$-5^{9}7^{3}13^{3}31 \cdot 47$
−8	−8	0
−11	−3	$-2^{12}11^{3}17^{3}29^{3}$
−11	−4	$-2^{12}7^{6}11 \cdot 19^{2}43^{2}$
−11	−7	$-7^{3}13^{3}17^{2}19^{2}41 \cdot 61 \cdot 73$
−11	−8	$-2^{12}7^{6}13^{3}29^{2}79$
−11	−11	$-2^{12}7^{6}11 \cdot 13^{3}17^{2}19^{2}$
−12	−3	0
−12	−4	$2^{10}3^{9}11^{2}23^{2}47^{2}$
−12	−7	$-3^{9}5^{9}17^{2}59 \cdot 83$
−12	−8	$-2^{10}5^{9}23 \cdot 29^{2}47 \cdot 71$
−12	−11	$-2^{19}11 \cdot 17^{2}41^{2}83 \cdot 107 \cdot 131$
−12	−12	$-2^{8}3^{9}5^{9}11^{2}17^{2}23$
−16	−3	$2^{9}3^{9}23^{3}47^{3}$
−16	−4	0
−16	−7	$-3^{20}7^{3}19^{2}31 \cdot 103$
−16	−8	$-2^{10}7^{6}23 \cdot 31 \cdot 47 \cdot 79 \cdot 103 \cdot 127$
−16	−11	$-2^{9}7^{6}11^{3}19^{2}127 \cdot 151 \cdot 167$
−16	−12	$-2^{7}3^{9}11^{6}23 \cdot 71 \cdot 167 \cdot 191$
−16	−16	$-2^{6}3^{20}7^{6}19^{2}23 \cdot 31$
−19	−3	$-2^{12}3^{9}41^{3}53^{3}$
−19	−4	$-2^{12}3^{20}19 \cdot 67^{2}$
−19	−7	$-3^{20}13^{3}19 \cdot 31^{2}97$
−19	−8	$-2^{12}13^{3}29 \cdot 31 \cdot 37 \cdot 71 \cdot 103 \cdot 127 \cdot 151$
−19	−11	$-2^{12}13^{3}19 \cdot 29^{2}41^{2}109 \cdot 173 \cdot 193$
−19	−12	$-2^{19}3^{9}29^{2}59 \cdot 107 \cdot 179 \cdot 227$
−19	−16	$-2^{9}3^{20}19 \cdot 31 \cdot 59^{2}79 \cdot 223$
−19	−19	$-2^{12}3^{20}13^{3}19 \cdot 29 \cdot 31^{2}37$

The body of this paper is divided into two parts (§§ 2—4 and §§ 5—7), as we have two proofs of the above results which are of an essentially different nature. The first method is algebraic, and works at the "finite primes". It relies on the work of Deuring on the endomorphism rings of elliptic curves, and exploits the connection between the arithmetic of maximal orders in quaternion algebras of prime discriminant l over $\mathbb{Q}$ and the geometry of supersingular elliptic curves in characteristic l. Some of these methods were already used by Deuring in [6]. The second method is analytic, and works at the "infinite primes". It is based on the calculation of the Fourier coefficients

of the restriction to the diagonal $\mathfrak{H} \subset \mathfrak{H} \times \mathfrak{H}$ of an Eisenstein series for the Hilbert modular group of $\mathbb{Q}(\sqrt{D})$, and was suggested by a paper of Siegel [18]. Both methods may be viewed as the special case $N=1$ of the theory of local heights of Heeger points on $X_0(N)$. The general case, and its relation to the derivatives of L-series and to forms of half-integral weight, will be treated in forthcoming papers [9], [10].

2. Let W be a complete, discrete valuation ring whose quotient field has characteristic zero and whose residue field is algebraically closed of characteristic $l > 0$. Let π be a prime of W, and normalize the valuation v so that $v(\pi) = 1$. We will adopt the convention that $v(0) = +\infty$.

Let E and E' be elliptic over W with good reduction (mod π). These curves have plane cubic models of the form

$$E: \quad y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6,$$
$$E': \quad y^2 + a_1' xy + a_3' y = x^3 + a_2' x^2 + a_4' x + a_6'$$

as in Tate [20], which will be our basic reference in this section. The coefficients a_i, a_i' are elements of W, and the discriminants Δ, Δ' are elements of W^*. Let $j = j(E)$ and $j' = j(E')$ be the modular invariants of the two curves; these are independent of the model chosen, and we have the identity

$$j - j' = \frac{c_6^2 c_4'^3 - c_6'^2 c_4^3}{1728 \Delta \Delta'} \quad \text{in } W. \tag{2.1}$$

For each integer $n \geqq 1$, the set $\mathrm{Iso}_n(E, E')$ of isomorphisms from E to E' which are defined over the ring W/π^n is finite of order 0, 2, 4, 6, 12, or 24. We define

$$i(n) = \frac{\mathrm{Card}\,(\mathrm{Iso}_n(E, E'))}{2}. \tag{2.2}$$

The main result of this section is the following

Proposition 2.3. $$v(j - j') = \sum_{n \geqq 1} i(n).$$

Note that this formula refines the well-known result that $v(j - j') > 0$ if and only if the curves E and E' are isomorphic over the algebraically closed field W/π. For the rest of this section, we shall assume that an isomorphism exists (mod π); otherwise, both sides of 2.3 are equal to zero.

Proof. We first assume that $l > 3$, so $v(1728) = 0$. Since $1728 \Delta = c_4^3 - c_6^2$, at least one of the quantities c_4, c_6 must be a unit in W. We may choose models for E and E' with $a_1 = a_1' = 0$, $a_2 = a_2' = 0$, and $a_3 = a_3' = 0$ in W. Then $c_4 = -2^4 3 a_4$ and $c_6 = -2^5 3^3 a_6$; hence one of the coefficients a_4, a_6 is a unit in W.

Since the curves E and E' are isomorphic (mod π), we can solve the simultaneous congruences:

$$\begin{cases} a_4 \equiv u^4 a_4' \\ a_6 \equiv u^6 a_6' \end{cases} \pmod{\pi}$$

for a unit $u \in (W/\pi)^*$. We have $i(n) \geqq 1$ if and only if these congruences can be solved (mod π^n).

Assume that $i(n) \geqq 1$. If a_4 is a unit, we may normalize $a_4 = a_4' = 1$ and choose (a_6, a_6') so that $v(a_6 - a_6')$ is maximal. When $i(n) \geqq 1$,

$$i(n) = \begin{cases} 2 & \text{if } a_6 \equiv a_6' \equiv 0 \pmod{\pi^n}, \\ 1 & \text{if } a_6 \not\equiv 0,\ a_6' \not\equiv 0 \pmod{\pi^n}. \end{cases}$$

Hence, $v(j-j') = v(a_6^2 - a_6'^2) = v(a_6 - a_6') + v(a_6 + a_6') = \sum i(n)$ as claimed.

If a_6 is a unit, we may normalize $a_6 = a_6'$ and modify (a_4, a_4') by a cube root of unity so that $v(a_4 - a_4')$ is maximal. When $i(n) \geqq 1$,

$$i(n) = \begin{cases} 3 & \text{if } a_4 \equiv a_4' \equiv 0 \pmod{\pi^n}, \\ 1 & \text{if } a_4 \not\equiv 0,\ a_4' \not\equiv 0 \pmod{\pi^n}. \end{cases}$$

Hence $v(j-j') = v(a_4^3 - a_4'^3) = v(a_4 - a_4') + v(a_4 - \zeta a_4') + v(a_4 - \zeta^2 a_4')$, where ζ is a primitive cube root of unity in W^*. Hence $v(j-j') = \sum i(n)$ as claimed.

Now assume $l = 3$. The proof then splits into several cases; we will treat the case when $b_2 \equiv b_2' \equiv 0 \pmod 3$, which is most useful in our applications, and leave the others to the reader. Changing models, we may assume $a_1 = a_1' = 0$, $a_2 = a_2' = 0$, $a_3 = a_3' = 0$ in W. The quantities a_4 and a_4' must then be units in W^*, and we may change models to insure $a_4 = a_4' = 1$ and $v(a_6 - a_6')$ is maximal. Then

$$v(j-j') = v\left(3^6 (a_6^2 - a_6'^2)\right) = 6v(3) + v(a_6 - a_6') + v(a_6 + a_6').$$

The curves E and E' are isomorphic $(\bmod \pi^n)$ if and only if the simultaneous congruences

$$\begin{cases} 3r \equiv 0 \\ u^4 \equiv 1 \\ u^6 a_6' \equiv a_6 + r + r^3 \end{cases} \pmod{\pi^n}$$

can be solved for $u \in (W/\pi^n)^*$ and $r \in W/\pi^n$. If $i(n) \geqq 1$ then

$$i(n) = \begin{cases} 6 & \text{if } 3 \equiv 0 \pmod{\pi^n}, \\ 2 & \text{if } 3 \not\equiv 0,\ a_6 \equiv a_6' \equiv 0 \pmod{\pi^n}, \\ 1 & \text{if } 3 \not\equiv 0,\ a_6 \not\equiv 0,\ a_6' \not\equiv 0 \pmod{\pi^n}. \end{cases}$$

Hence $v(j-j') = \sum i(n)$ as claimed.

Finally, assume $l = 2$. Again the proof breaks into several cases; we will only treat the case when $a_1 \equiv a_1' \equiv 0 \pmod 2$ here. Changing models, we may assume that

$$a_1 = a_1' = 0,\ a_2 = a_2' = 0, \text{ and } a_6 = a_6' = 0.$$

Then a_3 and a_3' are units; we may change models to insure that $a_3 = a_3' = 1$ and $v(a_4 - a_4')$ is maximal. Then

$$v(j-j') = v\left(2^{12} (a_4^3 - a_4'^3)\right) = 12v(2) + v(a_4 - a_4') + v(a_4 - \zeta a_4') + v(a_4 - \zeta^2 a_4').$$

The curves E and E' are isomorphic (mod π^n) if and only if the simultaneous congruences

$$\begin{cases} 2s \equiv 0 \\ 3r \equiv s^2 \\ u^3 \equiv 1 + 2t & (\text{mod } \pi^n) \\ u^4 a_4' \equiv a_4 - s + 3r^2 \\ t^2 + t \equiv r a_4 + r^3 \end{cases}$$

can be solved for $u \in (W/\pi^n)^*$ and r, s, t in W/π^n. If $i(n) \geqq 1$, then

$$i(n) = \begin{cases} 12 & \text{if } 2 \equiv 0 & & (\text{mod } \pi^n), \\ 3 & \text{if } 2 \not\equiv 0, & a_4 \equiv a_4' \equiv 0 & (\text{mod } \pi^n), \\ 1 & \text{if } 2 \not\equiv 0, & a_4 \not\equiv 0, \ a_4' \not\equiv 0 & (\text{mod } \pi^n). \end{cases}$$

Hence $v(j-j') = \sum i(n)$ as claimed.

Corollary 2.4. *If $(l-1)$ divides 12 and the curves E and E' both have supersingular reduction* (mod π), *then* $v(j-j') \geqq \dfrac{12}{l-1}$.

Proof. In this case, there is a single supersingular invariant in characteristic l, so $E \cong E'$ over W/π. But then $v(j-j') \geqq i(1) = \dfrac{12}{l-1}$.

We are in a position to prove Berwick's congruences ([2], pg. 66—76) for the moduli $j = j(\tau)$ of an imaginary quadratic argument of discriminant d. Put $\mathcal{O} = \mathbb{Z}[\tau]$ and $K = \mathbb{Q}(\tau) = \mathbb{Q}(\sqrt{d})$.

Corollary 2.5. 1) *If $d < -4$ and $\left(\dfrac{d}{l}\right) = 1$ then*

$$\mathrm{N}(j)\,\mathrm{N}(j-1728) \not\equiv 0 \ (\text{mod } l).$$

2) *If $\left(\dfrac{d}{l}\right) = -1$ and $l < 12$ we have:*

$$\begin{array}{lll} j \equiv 0 & (\text{mod } 2^{15}) & l = 2, \\ j \equiv 1728 & (\text{mod } 3^6) & l = 3, \\ j \equiv 0 & (\text{mod } 5^3) & l = 5, \\ j \equiv 1728 & (\text{mod } 7^2) & l = 7, \\ j^{\frac{1}{3}}(j-1728)^{\frac{1}{2}} \equiv 0 & (\text{mod } 11) & l = 11. \end{array}$$

Proof. 1) If $\left(\dfrac{d}{l}\right) = 1$ then the elliptic curve E with invariant j has ordinary reduction (mod π) for all primes dividing l in $K(j)$. Furthermore, by Deuring's theory [5], $\mathrm{End}_{W/\pi}(E) = \mathrm{End}_W(E)$ is the ring $\mathcal{O}$ of complex multiplications. When $d < -4$ we have $\mathcal{O}^\times = \{\pm 1\}$; hence $j \not\equiv 0$ (mod π) and $j \not\equiv 1728$ (mod π).

2) If $\left(\frac{d}{l}\right) = -1$ then E has supersingular reduction (mod π) for all primes dividing l in $K(j)$ [5]. When $l=3$, $j'=1728$ is the unique supersingular invariant; when $l=5$, $j'=0$ is the unique supersingular invariant; when $l=7$, $j'=1728$ is the unique supersingular invariant; and when $l=11$, $j'=0$, 1728 are the unique supersingular invariants. This gives the congruences for $l \neq 2$, using 2. 4.

When $l=2$, $j=0$ is the unique supersingular invariant, so 2. 4 gives $j \equiv 0(2^{12})$, using the fact that $i(1)=12$ in 2. 3. To obtain the full congruence, we will show that $i(2)=3$. The ring $\mathrm{End}_{W/\pi}(E)$ is isomorphic to Hurwitz's order $\mathbb{Z}\left[i, j, k, \frac{1+i+j+k}{2}\right]$ in Hamilton's quaternions, and the subring $\mathrm{End}_{W/\pi^2}(E)$ has index 4 and contains $2\,\mathrm{End}_{W/\pi}(E)$ (Gross). The elements of order 6 in $\mathrm{End}_{W/\pi}(E)$ have the form $\frac{1 \pm i \pm j \pm k}{2}$, and one of these will be contained in End_{W/π^2} *unless* all elements in that ring have even reduced trace. But $\mathrm{End}_{W/\pi^2}(E)$ contains $\mathcal{O}$, which has elements of odd trace. Hence $i(2)=3$ and $j \equiv 0 \pmod{2^{15}}$.

We now present a refinement of Deuring's lifting theorem [5]. Let E_0 be an elliptic curve over the ring W/π^n, and let α_0 denote an endomorphism of E_0. *Assume* that the subring $\mathbb{Z}[\alpha_0] \subseteq \mathrm{End}_{W/\pi^n}(E_0)$ has rank 2 as a $\mathbb{Z}$-module and is integrally closed in its quotient field. Another way to express this is to associate to the endomorphism α_0 its trace $t=\alpha_0+\alpha_0^\vee$ and norm $n=\alpha_0 \circ \alpha_0^\vee$ which may be viewed as multiplication by fixed integers in $\mathrm{End}_{W/\pi^n}(E_0)$. Our assumption is then that the integer $d=t^2-4n$ is a fundamental negative discriminant.

On the tangent space $\mathrm{Lie}(E_0)$, α_0 induces multiplication by an element w_0 which satisfies the quadratic equation $x^2-tx+n=0$. Clearly, a necessary condition for lifting E_0 with the endomorphism α_0 to W is the existence of an element $w \in W$ which satisfies

$$\begin{cases} w \equiv w_0 \pmod{\pi^n}, \\ w^2 - tw + n = 0, \end{cases} \tag{2.6}$$

as the induced action of the lifted endomorphism on the tangent space will give rise to such an element.

Proposition 2. 7. *Suppose a w exists which satisfies* (2. 6). *Then there is an elliptic curve E over W and an endomorphism* α *of E such that*

a) $(E, \alpha) \equiv (E_0, \alpha_0) \bmod \pi^n$;

b) α *induces multiplication by w on* $\mathrm{Lie}(E)$.

If (E', α') *is any other lifting, there is a commutative diagram*

$$\begin{array}{ccc} E & \xrightarrow{\alpha} & E \\ \downarrow \wr & & \downarrow \wr \\ E' & \xrightarrow{\alpha'} & E' \end{array}$$

of morphisms over W.

Proof. Let l be the characteristic of W/π. By the deformation theory of Serre and Tate [16], it suffices to construct a lifting of the l-divisible group $\hat{E}_0$ of E_0, together with an endomorphism lifting $\hat{\alpha}_0$.

When E_0 is ordinary over W/π, we may take E to be the canonical lifting. This is the unique curve reducing to E_0 where $\hat{E}$ is the direct product of a group of multiplicative type with an étale group. Since $\mathrm{End}_W(E) = \mathrm{End}_{W/\pi^n}(E_0) = \mathrm{End}_{W/\pi}(E_0)$, we clearly have an endomorphism lifting α.

When E_0 is supersingular over W/π, we may lift $\hat{E}_0$ to a Lubin-Tate group $\hat{E}$ of height 2 over W with endomorphism $\alpha[x] = wx + \cdots$. The uniqueness of this lifting shows that it is algebraic.

3. We now turn to the algebraic proof of 1. 3. We shall assume that $d_1 = -p$ is *prime*, which facilitates some of the computations, but the method works quite generally. (See Dorman [7] for the details.) For d_2 we take an arbitrary negative fundamental discriminant prime to p.

Let $\tau = \dfrac{1+\sqrt{-p}}{2}$, so $\mathcal{O} = \mathbb{Z}[\tau]$ is the ring of integers in $K = \mathbb{Q}(\sqrt{-p})$. Let $j = j(\tau)$; then $H = K(j)$ is the Hilbert class field of K. If v is a finite place of H, we let A_v denote the completion of the maximal unramified extension of the ring of v-integers in H, and let W_v denote the extension of A_v obtained by adjoining a *fixed* element w which satisfies an integral quadratic equation of discriminant d_2. This extension will be non-trivial if and only if the residual characteristic l of A_v divides d_2. We let e denote the ramification index of W_v/A_v.

Define the algebraic integer $\alpha = \alpha(\tau, d_2)$ by:

$$\alpha = \prod_{\substack{[\tau_2] \\ \mathrm{disc}\,\tau_2 = d_2}} (j - j(\tau_2))^{\frac{4}{w_1 w_2}}. \tag{3.1}$$

This lies in H, and when $d_2 \neq -4$ even lies in the subfield $\mathbb{Q}(j)$. Our aim is to calculate the valuation of α at each finite place v of H, using the methods of § 2. To do this, let E be an elliptic curve over $W = W_v$ with multiplication by $\mathcal{O}$ and invariant $j(E) = j$. This existence of such a curve with good reduction is guaranteed by a theorem of Serre and Tate [17]; it is unique up to W-isomorphism as the residue field is algebraically closed. Similarly, for each τ_2 of discriminant d_2, let E' denote an elliptic curve over W with multiplication by $\mathbb{Z}[w]$ and invariant $j' = j(\tau_2)$; then by 2. 3 we have

$$\mathrm{ord}_v(\alpha) = \frac{4}{e w_1 w_2} \sum_{\substack{[\tau_2] \\ \mathrm{disc}\,\tau_2 = d_2}} \sum_{n \geq 1} \frac{1}{2} \# \mathrm{Iso}_{W/\pi^n}(E, E'). \tag{3.2}$$

We are therefore reduced to counting isomorphisms $f: E \xrightarrow{\sim} E' \pmod{\pi^n}$. Such an isomorphism gives rise to an endomorphism $w_f = f^{-1} \cdot w \cdot f$ of $E \pmod{\pi^n}$ which belongs to the set

$$S_n = \{\alpha_0 \in \mathrm{End}_{W/\pi^n} E \mid \mathrm{Tr}(\alpha_0) = \mathrm{Tr}(w),\ \mathrm{N}(\alpha_0) = \mathrm{N}(w),\ \alpha_0 = w \text{ on } \mathrm{Lie}(E)\}.$$

Furthermore, proposition (2. 6) insures that every element α_0 of S_n is of the form w_f for some isomorphism $f: E \to E'$ (mod π^n) to a curve E' with complex multiplication by $\mathbb{Z}[w]$. Indeed, the pair (E, α_0) can be lifted to (F, α) over W and since F has multiplication by $\mathbb{Z}[\alpha] = \mathbb{Z}[w]$ it is isomorphic to one of the curves E' via a map $f: F \xrightarrow{\sim} E'$ with $\alpha = f^{-1} \cdot w \cdot f$. Reducing this map (mod π^n) shows that $\alpha_0 = w_f$; (2. 6) also gives the uniqueness of E' over W as well as the uniqueness of f up to a W-automorphism of E'. Hence,

$$\text{ord}_v(\alpha) = \frac{2}{e w_1} \sum_{n \geqq 1} \# S_n. \tag{3. 3}$$

We now turn to a computation of the set S_n. Recall that l is the residual characteristic of v.

Lemma 3. 4. *If* $\left(\frac{l}{p}\right) = 1$, *then* $\text{ord}_v(\alpha) = 0$.

Proof. In this case, E has ordinary reduction (mod π) and $\text{End}_{W/\pi^n} E = \mathcal{O}$ for all $n \geqq 1$ [5]. Since this ring contains no elements of discriminant d_2, $\# S_n = 0$ for all $n \geqq 1$.

Now suppose that $\left(\frac{l}{p}\right) \neq 1$; then E has supersingular reduction (mod π) and $\text{End}_{W/\pi} E$ is isomorphic to a maximal order in the quaternion algebra B over $\mathbb{Q}$ which is ramified at l and ∞ [5]. Our first task is a convenient description of this order, as well as its subrings $\text{End}_{W/\pi^n} E$ for $n \geqq 1$.

Since $\left(\frac{l}{p}\right) \neq 1$, the field $\mathbb{Q}(j)$ has a unique embedding into the field $\mathbb{Q}_l$ of l-adic numbers [8]. If v_1 is the place of $H = K(j)$ which corresponds to the two equivalent extensions of this embedding, there is a unique element σ in $G = \text{Gal}(H/K)$ such that $\text{ord}_v(\beta) = \text{ord}_{v_1}(\beta^\sigma)$ for all $\beta \in H^*$. Let $\mathfrak{a}$ be a fractional ideal in K whose class corresponds to σ under the Artin isomorphism. The algebra B is given by the subring $\left\{[\alpha, \beta] = \begin{pmatrix} \alpha & \beta \\ -l\bar{\beta} & \bar{\alpha} \end{pmatrix}\right\}$ of the 2×2 matrices over K; let $\mathcal{D}^{-1}$ denote the inverse different of $\mathcal{O}$ and λ a fixed solution of the congruence $\lambda^2 \equiv -l \pmod{\mathcal{D}}$.

From now on we write $d_2 = -q$, $q > 0$.

Lemma 3. 5. *Assume* $l \nmid pq$. *Then* $e = 1$ *and*

1) $\text{End}_{W/\pi^n} E = \{[\alpha, \beta] : \alpha \in \mathcal{D}^{-1}, \beta \in \mathcal{D}^{-1} l^{n-1} \bar{\mathfrak{a}}/\mathfrak{a}, \alpha \equiv \lambda\beta \bmod \mathcal{O}_p\}$.

2) *The number of elements of* S_n *is equal to* $\frac{w_1}{2}$ *times the number of solutions* $(x, \mathfrak{b})$ *of the equation* $x^2 + 4l^{2n-1} \text{N}\mathfrak{b} = pq$, *where* x *is an integer and* $\mathfrak{b}$ *is an ideal of* $\mathcal{O}$ *in the class of* $\mathfrak{a}^2$, *the solutions* $(x, \mathfrak{b})$ *with* $x \equiv 0 \pmod{p}$ *being counted twice.*

Proof. 1) When $\mathfrak{a} \sim 1$ the ring $\text{End}_{W/\pi} E$ contains $\text{End}_W E = \{[\alpha, 0] \alpha \in \mathcal{O}\}$ as well as the $\mathcal{O}$-span of the Frobenius endomorphism $F = \begin{pmatrix} 0 & 1 \\ -l & 0 \end{pmatrix} = [0,1]$, since the

reduced curve descends to the prime field of l elements. Since $\mathrm{End}_{W/\pi} E$ is a maximal order in B, it must be isomorphic to $\{[\alpha,\beta] | \alpha \in \mathscr{D}^{-1},\ \beta \in \mathscr{D}^{-1},\ \alpha \equiv \lambda\beta \bmod \mathcal{O}\}$. The calculation of End_{W/π^n} then follows from the observation that $l^k F$ is an endomorphism of $E(\mathrm{mod}\ \pi^n)$ if and only if $k \geqq n-1$.

When $\mathfrak{a}$ is arbitrary, $\mathrm{End}_{W/\pi} E^{\sigma_\mathfrak{a}}$ is isomorphic to the ring calculated above. Since $\mathrm{Hom}_W(E^{\sigma_\mathfrak{a}}, E)$ is isomorphic to $\mathfrak{a}$ as a left $\mathcal{O} = \mathrm{End}_W(E)$-module [3], we have $\mathfrak{a} \cdot \mathrm{End}_{W/\pi^n} E = \mathrm{End}_{W/\pi^n} E^{\sigma_\mathfrak{a}} \cdot \mathfrak{a}$ in B. A short calculation yields the desired result.

2) If $[\alpha, \beta]$ is an endomorphism of $E(\mathrm{mod}\ \pi^n)$ with trace $= \mathrm{Tr}(w)$ and norm $= \mathrm{N}(w)$ then $\alpha = \dfrac{x + \mathrm{Tr}(w)\sqrt{-p}}{2\sqrt{-p}}$ with x an integer and $\beta = \dfrac{\gamma l^{n-1}}{\sqrt{-p}}$ with $\gamma \in \bar{\mathfrak{a}}/\mathfrak{a}$.

Let $\mathfrak{b} = (\gamma)\,\mathfrak{a}/\bar{\mathfrak{a}}$; then $\mathfrak{b}$ is an integral ideal in the class of $\mathfrak{a}^2$ and

$$x^2 + 4l^{2n-1}\mathrm{N}\mathfrak{b} = pq.$$

Conversely if $(x, \mathfrak{b})$ solves this equation with $\mathfrak{b}$ in the class of $\mathfrak{a}^2$, and γ is any generator of the principal ideal $\mathfrak{b}\bar{\mathfrak{a}}/\mathfrak{a}$, we may obtain an element $[\alpha, \beta]$ of B with trace $= \mathrm{Tr}(w)$ and norm $= \mathrm{N}(w)$ by reversing the above definitions. To determine whether or not $[\alpha, \beta]$ lies in $\mathrm{End}_{W/\pi^n} E$ we must test the congruence $\alpha \overset{?}{\equiv} \lambda\beta \bmod \mathcal{O}_p$. This will hold for $\frac{w_1}{2}$ choices of the generator γ if $x \not\equiv 0 \pmod p$, and for w_1 choices if $x \equiv 0 \pmod p$. To see this, note that for any choice of generator we have $\alpha^2 \equiv -l\beta^2 \bmod \mathcal{O}_p$. For this is equivalent to the congruence $x^2 \equiv -4l^{2n-1}\gamma^2 \pmod{\sqrt{-p}}$, which follows from the fact that

$$\gamma^2 \equiv \gamma\bar{\gamma} = \mathrm{N}\gamma = \mathrm{N}\mathfrak{b} \pmod{\sqrt{-p}}.$$

Finally, to count the elements of S_n, we must determine which endomorphisms $[\alpha, \beta]$ induce multiplication by w on $\mathrm{Lie}(E)$. If $[\alpha, \beta]$ has this property, then the dual endomorphism $[\alpha, \beta]^\vee = [\bar{\alpha}, -\beta]$ induces multiplication by $\bar{w} \not\equiv w (\mathrm{mod}\ \pi)$. Hence we may count elements of S_n by taking exactly one half of the solutions $(x, \mathfrak{b})$.

Now consider the case when $l | q$. Lemma 3. 5 gives the endomorphism ring of E over $A_v/l^n A_v$, and W is a quadratic ramified extension of A_v. We therefore find

$$\mathrm{End}_{W/\pi^n} E = \{[\alpha, \beta] | \alpha \in \mathscr{D}^{-1},\ \beta \in l^{m-1}\,\bar{\mathfrak{a}}/\mathfrak{a},\ \alpha \equiv \beta (\mathrm{mod}\ \mathcal{O}_p)\} \quad \text{with} \quad m = \left[\frac{n+1}{2}\right].$$

The elements α_0 of this ring of trace $= \mathrm{Tr}(w)$ and norm $= \mathrm{N}(w)$ give solutions $(x, \mathfrak{b})$ of the equation $x^2 + 4l^{2m-1}\mathrm{N}\mathfrak{b} = pq$ as in 3. 5. Clearly such solutions can exist only when $m = 1$, so $n \leqq 2$. Since α_0 induces multiplication by an element of W/π on $\mathrm{Lie}(E)$, and the reduction of $w (\mathrm{mod}\ \pi^2)$ does not lie in the residue field, we see that S_n is empty for $n \geqq 2$. Since $w \equiv \bar{w} (\mathrm{mod}\ \pi)$ we have the equality:

$$S_1 = \{\alpha_0 \in \mathrm{End}_{W/\pi} E : \mathrm{Tr}(\alpha_0) = \mathrm{Tr}(w) \text{ and } \mathrm{N}(\alpha_0) = \mathrm{N}(w)\}.$$

Hence

Lemma 3. 6. *Assume $l | q$. Then $e = 2$ and S_n is empty for $n \geqq 2$. The number of elements in S_1 is $\frac{w_1}{2}$ times the number of solutions of the equation $x^2 + 4l\mathrm{N}\mathfrak{b} = pq$, where x is an integer and $\mathfrak{b}$ an ideal of $\mathcal{O}$ in the class of $\mathfrak{a}^2$, the solutions $(x, \mathfrak{b})$ with $x \equiv 0$ (mod p) being counted twice.*

Finally, we turn to the case where $l=p$. A computation similar to 3. 5 gives the result: $\mathrm{End}_{W/\pi^n} E=\{[\alpha,\beta]:\alpha\in\mathcal{O},\ \beta\in\mathcal{D}^{n-2}\bar{\mathfrak{a}}/\mathfrak{a}\}$. Let $\alpha_0=[\alpha,\beta]$ be an element of trace $=\mathrm{Tr}(w)$ and norm $=\mathrm{N}(w)$ and write $\alpha=\dfrac{\mathrm{Tr}(w)+y\sqrt{-p}}{2}$, $\beta=\dfrac{\gamma(\sqrt{-p})^{n-1}}{\sqrt{-p}}$ with y an integer and $\gamma\in\bar{\mathfrak{a}}/\mathfrak{a}$. Letting $x=py$ and $\mathfrak{b}=(\gamma)\,\bar{\mathfrak{a}}/\mathfrak{a}$ we find a solution to the equation $x^2+4p^n\mathrm{N}\mathfrak{b}=pq$. Hence $n=1$. Conversely, any solution with $\mathfrak{b}$ integral in the class of $\mathfrak{a}^2$ gives exactly w_1 elements α_0 with the correct trace and norm, as this is the number of choices for the generator γ. Again, exactly half of these elements will lie in S_1, as $w\not\equiv\bar{w}\,(\mathrm{mod}\ \pi)$, so we have

Lemma 3. 7. *Assume $l=p$. Then $e=1$ and S_n is empty for $n\geqq 2$. The number of elements in S_1 is $\frac{w_1}{2}$ times the number of solutions $(x,\mathfrak{b})$ of the equation $x^2+4p\,\mathrm{N}\mathfrak{b}=pq$, where x is an integer (divisible by p) and $\mathfrak{b}$ is an ideal of $\mathcal{O}$ in the class of $\mathfrak{a}^2$.*

We may combine the last three lemmas as follows. For $m\geqq 1$ we let $r_{\mathfrak{a}^2}(m)$ denote the number of ideals of $\mathcal{O}$ in the class of $\mathfrak{a}^2$ of norm equal to m. For an integer x, we define

$$\delta(x)=\begin{cases}2 & \text{if } x\equiv 0\,(\mathrm{mod}\ p),\\ 1 & \text{otherwise.}\end{cases}$$

Proposition 3. 8. *Assume $\left(\frac{l}{p}\right)\neq 1$. Let v be a finite place dividing l and $\mathfrak{a}$ the ideal defined before* 3. 5. *Then*

$$\mathrm{ord}_v(\alpha)=\frac{1}{2}\sum_{n\in Z}\sum_{n\geqq 1}\delta(x)\,r_{\mathfrak{a}^2}\left(\frac{pq-x^2}{4l^n}\right).$$

Next note that by (1. 2) and (3. 1) we have the relation:

$$J(-p,-q)=\mathrm{N}_{H/K}(\alpha). \tag{3. 9}$$

Furthermore, the sum $\sum_{\mathfrak{a}} r_{\mathfrak{a}^2}(m)$ is equal to the number $R(m)$ of ideals of $\mathcal{O}$ of norm m, as the class group of $\mathbb{Q}(\sqrt{-p})$ has *odd* order. Hence we have

Proposition 3. 9. *If λ is a prime of $\mathcal{O}$ of characteristic l, then*

$$\mathrm{ord}_\lambda J(-p,-q)=\frac{1}{2}\sum_{x\in Z}\sum_{n\geqq 1}\delta(x)\,R\left(\frac{pq-x^2}{4l^n}\right).$$

It is an exercise to derive theorem 7. 3 from this proposition, using the identity $R(m)=\sum_{\substack{n\mid m\\ n>0}}\left(\frac{n}{p}\right)$ afforded by Dirichlet's factorization of the zeta-function of K.

4. To express the results in the previous section neatly, and to find appropriate generalizations, it is convenient to introduce the modular polynomials. For each negative discriminant d, we define

$$f_d(x)=\prod_{f^2\mid d}\prod_{\substack{[\tau]\\ \mathrm{disc}\,\tau=\frac{d}{f^2}}}(x-j(\tau))^{2/w(d/f^2)}. \tag{4. 1}$$

This "polynomial" has integral coefficients and degree

$$H(|d|)=\sum_{f^2|d}\frac{2h\left(\frac{d}{f^2}\right)}{w\left(\frac{d}{f^2}\right)},$$

the Hurwitz class number. The first few examples are

$$\begin{aligned}
f_{-3}(x) &= x^{\frac{1}{3}}, & H(3) &= \frac{1}{3},\\
f_{-4}(x) &= (x-1728)^{\frac{1}{2}}, & H(4) &= \frac{1}{2},\\
f_{-7}(x) &= x+3375, & H(7) &= 1,\\
f_{-8}(x) &= x-8000, & H(8) &= 1,\\
f_{-11}(x) &= x+32768, & H(11) &= 1,\\
f_{-12}(x) &= x^{\frac{1}{3}}(x-54000), & H(12) &= \frac{4}{3}.
\end{aligned}$$

In the last section, we factored $\alpha = f_{d_2}(j_1)^{\frac{2}{w_1}}$ in the integers of the field $\mathbb{Q}(j_1)$. Note that when $(d_1, d_2)=1$, the value $J(d_1, d_2)$ is equal to the resultant of the polynomials $f_{d_1}(x)$ and $f_{d_2}(x)$.

For $m \geqq 1$, let $\varphi_m(x, y)$ be the polynomial in $\mathbb{Z}[x, y]$ defined by

$$\varphi_m(j(z_1), j(z_2)) = \prod_{\substack{\det\gamma = m \\ \bmod SL_2(\mathbb{Z})}} (j(z_1) - j(\gamma z_2)). \tag{4.2}$$

Here the product is taken over the equivalence classes of 2×2 integral matrices of determinant m, modulo the left action of $SL_2(\mathbb{Z})$. The polynomial $\varphi_m(x, y)$ is often referred to as the "modular equation of level m", although the usual definition takes the product only over the primitive classes γ in order to obtain an *irreducible* curve in $\mathbb{P}^1 \times \mathbb{P}^1$ which is a model for $X_0(m)$. The fact that $\varphi_m(x, y)$ has integral coefficients is well-known. We have

$$\varphi_1(x, y) = x - y,$$

$$\varphi_2(x, y) = x^3 + y^3 - x^2y^2 + 2^4\cdot 3\cdot 31(x^2y + y^2x) - 2^4\cdot 3^4\cdot 5^3(x^2+y^2) + 3^4\cdot 5^3\cdot 4027\,xy + 2^8\cdot 3^7\cdot 5^6(x+y) - 2^{12}\cdot 3^9\cdot 5^9;$$

for the tabulation of φ_3, φ_5, and φ_7 see [14].

The polynomial $\varphi_m(x, y)$, when restricted to the diagonal, is related to the polynomials $f_d(x)$ by Kronecker's identity

$$\varphi_m(x, x) = \pm \prod_{\substack{t\in\mathbb{Z} \\ t^2<4m}} f_{t^2-4m}(x), \tag{4.3}$$

which holds whenever m is not a perfect square. Taking the degrees of both sides of (4. 3) gives the famous Kronecker-Hurwitz class number relation

$$\sum_{\substack{m=dd' \\ d>0}} \max(d, d') = \sum_{\substack{t^2<4m \\ t\in\mathbb{Z}}} H(4m-t^2), \tag{4. 4}$$

which is the weight 2 case of the Eichler-Selberg trace formula on $PSL_2(\mathbb{Z})$. The identity (4. 3) can be extended to hold for all m, provided we replace the term $(x-y)$ which divides $\varphi_m(x, y)$ when m is a square, by $\prod_{t^2<4} f_{t^2-4}(x) = x^{\frac{2}{3}}(x-1728)^{\frac{1}{2}}$ in the limit. Similarly, (4. 4) holds for all m if we take the sum over $t^2 \leqq 4m$ and define

$$H(0) = \zeta(-1) = -\frac{1}{12}.$$

Now suppose $j = j\left(\frac{1+\sqrt{-p}}{2}\right)$ is a singular modulus of discriminant $-p$, and that $m \geqq 1$ is *not* the norm of an element $\frac{a+b\sqrt{-p}}{2}$ in $\mathcal{O}$. Then the value $\Phi_m(j, j)$ is non-zero; by 3. 8 and (4. 3) we have the formula

$$\begin{aligned} \operatorname{ord}_v\left(\varphi_m(j, j)^{\frac{4}{w^2}}\right) &= \frac{2}{w} \sum_{t^2<4m} \left\{ \frac{1}{2} \sum_x \sum_{k\geqq 1} \delta(x) r_{a^2}\left(\frac{p(4m-t^2)-x^2}{4l^k}\right)\right\} \\ &= \sum_{n\geqq 0} \sum_{k\geqq 1} \delta(n)\, r_1(n)\, r_{a^2}\left(\frac{mp-n}{l^k}\right), \end{aligned} \tag{4. 5}$$

where $r_1(n)$ counts representations of n as $\frac{x^2+pt^2}{4}$ (i.e. as the norm of an ideal in the principal class) and we define $r_{\mathfrak{b}}(0) = \frac{1}{w}$ for any class $\mathfrak{b}$. We can generalize (4. 5) as follows. Let $\mathfrak{b}$ be an ideal of $\mathcal{O}$ and $m \geqq 1$ an integer which is *not* the norm of an ideal in the class of $\mathfrak{b}$. Then the element

$$\beta = \varphi_m(j, j^{\sigma_{\mathfrak{b}}})^{\frac{4}{w^2}} \tag{4. 6}$$

is a non-zero algebraic integer, and the following result gives its valuation at places v of H.

Theorem 4. 7. *If* $\left(\frac{l}{p}\right) = 1$ *then* $\operatorname{ord}_v(\beta) = 0$. *If* $\left(\frac{l}{p}\right) \neq 1$ *and we define* $\mathfrak{a}$ *as in Lemma* 3. 5, *then*

$$\operatorname{ord}_v(\beta) = \sum_{n\geqq 0} \sum_{k\geqq 1} \delta(n)\, r_{\mathfrak{b}^{-1}}(n)\, r_{\mathfrak{b}\mathfrak{a}^2}\left(\frac{mp-n}{l^k}\right).$$

The case when $m = 1$ and $\mathfrak{b}$ is not principal is particularly interesting, as (4. 7) gives the prime factorization of $(j-j^{\sigma_{\mathfrak{b}}})$. Taking the norm of this quantity to K and then the product over all classes $\mathfrak{b} \neq 1$ gives the discriminant of the monic polynomial

of degree h satisfied by j. From this, we can obtain a formula for the index I of the order $\mathbb{Z}[j]$ in the ring of integers of $\mathbb{Q}(j)$, as the field discriminant is equal to $(-p)^{\frac{h-1}{2}}$ [8].

Corollary 4.8. *For any rational prime l we have*

$$\mathrm{ord}_l(I) = \sum_{n\geq 1}\sum_{k\geq 1} \frac{R(n)-r_1(n)}{2} R\left(\frac{p-n}{l^k}\right).$$

In particular, if l divides I then $l<p$ and $\left(\frac{l}{p}\right)=-1$.

We now sketch the proof of 4.7, again using the methods of § 2. We will assume for simplicity that $\left(\frac{l}{p}\right)=-1$ and that l does not divide m. Let W be the completion of the maximal unramified extension of the ring of v-integers of H, and let E be an elliptic curve over W with multiplication by $\mathcal{O}$ and invariant j as in § 3. By (3.5) we have

$$(4.9)\quad \mathrm{End}_{W/\pi^n} E = \{[\alpha,\beta] : \alpha\in\mathscr{D}^{-1},\ \beta\in\mathscr{D}^{-1}l^{n-1}\bar{\mathfrak{a}}/\mathfrak{a},\ \alpha\equiv\lambda\beta \bmod \mathcal{O}_p\}.$$

Let $\mathfrak{b}$ be an ideal of $\mathcal{O}$ which is prime to l and in the class of σ. By [17] we have $\mathrm{Hom}_W(E^{\sigma_\mathfrak{b}}, E)\cong\mathfrak{b}$ as an $\mathcal{O}=\mathrm{End}_W(E)$-module. Hence $\mathrm{Hom}_{W/\pi^n}(E^{\sigma_\mathfrak{b}}, E)=\mathrm{End}_{W/\pi^n}E\cdot\mathfrak{b}$ inside $B_{l,\infty}$. Hence

$$(4.10)\quad \mathrm{Hom}_{W/\pi^n}(E^{\sigma_\mathfrak{b}}, E) = \{[\alpha,\beta] : \alpha\in\mathscr{D}^{-1}\mathfrak{b},\ \beta\in\mathscr{D}^{-1}l^{n-1}\bar{\mathfrak{b}}\bar{\mathfrak{a}}/\mathfrak{a},\ \alpha\equiv\lambda\beta \bmod \mathcal{O}_p\}.$$

If $\varphi=[\alpha,\beta]$ has degree m, then $\mathrm{N}\alpha+l\mathrm{N}\beta=m\mathrm{N}\mathfrak{b}$.

On the other hand, by the definition of φ_m and the results in § 2, we have

$$(4.11)\qquad \mathrm{ord}_v(\beta)=\frac{4}{w^2}\sum_{n\geq 1}\frac{\mathrm{Card}\left(\mathrm{Hom}_{W/\pi^n}(E^{\sigma_\mathfrak{b}}, E)_{\mathrm{degree}\,m}\right)}{2}.$$

If $\alpha=\dfrac{\gamma}{\sqrt{-p}}$ and $\beta=\dfrac{l^{n-1}\delta}{\sqrt{-p}}$ correspond to an isogeny of degree m we have

$$(4.12)\qquad \mathrm{N}\mathfrak{c}+l^{2n-1}\mathrm{N}\mathfrak{d}=mp$$

where $\mathfrak{c}=(\gamma)/\mathfrak{b}$ and $\mathfrak{d}=(\delta)\,\mathfrak{a}/\bar{\mathfrak{b}}\bar{\mathfrak{a}}$ are integral ideals of $\mathcal{O}$ in the classes of σ^{-1} and $\sigma\sigma_{\mathfrak{a}^2}$ respectively. Conversely, given a solution to (4.12) we retrieve either $2\cdot\left(\frac{w}{2}\right)^2$ or $2w^2$ elements $[\alpha,\beta]$ of degree m in $\mathrm{Hom}_{W/\pi^n}(E^\sigma, E)$ by choosing generators for the principal ideals $\mathfrak{b}\mathfrak{c}$ and $\bar{\mathfrak{b}}\bar{\mathfrak{a}}/\mathfrak{a}\mathfrak{d}$. The second case occurs when $\mathrm{N}\mathfrak{c}\equiv 0(p)$. This completes our sketch of the proof of 4.7.

5. The analytic approach to the theorems of this paper consists of two parts: first, to give an expression for $\log|j(\tau_1)-j(\tau_2)|$ as an infinite sum over $PSL_2(\mathbb{Z})$ (or, rather, as a limit of such sums) which for imaginary quadratic arguments can be rewritten as a sum over rational integers, and secondly, to show that certain combinations of these infinite sums equal finite sums of logarithms of rational numbers. We carry out the first part in this section.

For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$ let Q_{s-1} be the Legendre function of the second kind, defined by

$$Q_{s-1}(t) = \int_0^\infty (t + \sqrt{t^2-1}\cosh v)^{-s}\, dv \quad (t > 1)$$

or

$$Q_{s-1}\left(\frac{1+t}{1-t}\right) = \frac{\Gamma(s)^2}{2\Gamma(2s)}(1-t)^s F(s, s; 2s; 1-t) \quad (0 < t < 1)$$

([1], 3.2 (36); F = hypergeometric function), and define for $\tau_j = u_j + iv_j \in \mathfrak{H}$ $(j = 1,2)$

$$g_s(\tau_1, \tau_2) = -2Q_{s-1}(\cosh d(\tau_1, \tau_2)) = -2Q_{s-1}\left(\frac{(u_1-u_2)^2 + v_1^2 + v_2^2}{2v_1 v_2}\right),$$

where d denotes hyperbolic distance. This is not defined at $\tau_1 = \tau_2$ since g_s has a singularity $\log|\tau_1 - \tau_2|^2$ along the diagonal. Because $d(\tau_1, \tau_2) = d(\gamma\tau_1, \gamma\tau_2)$ for any $\gamma \in PSL_2(\mathbb{R})$, the function G_s defined by the absolutely convergent series

$$G_s(\tau_1, \tau_2) = \sum_{\gamma \in \Gamma} g_s(\tau_1, \gamma\tau_2) \quad (\Gamma = PSL_2(\mathbb{Z}))$$

is Γ-invariant in each variable separately. The function G_s is called the *automorphic Green's function* or *resolvent kernel* and is studied in various places, e.g. [13] (note that our function is 4π times Hejhal's). The properties we need are:

a) G_s is real-analytic on $(\Gamma\backslash\mathfrak{H})^2 \setminus$(diagonal) but has a singularity $\log|\tau_1 - \tau_2|^2 + O(1)$ as $\tau_2 \to \tau_1$;

b) $\Delta_j G_s = s(s-1)\, G_s$, where Δ_j $(j = 1, 2)$ is the hyperbolic Laplace operator $v_j^2\left(\frac{\partial^2}{\partial u_j^2} + \frac{\partial^2}{\partial v_j^2}\right)$;

c) For τ_2 fixed and $v_1 = \mathrm{Im}(\tau_1)$ large (larger than $\max_{\gamma \in \Gamma} \mathrm{Im}(\gamma\tau_2)$), G_s has a Fourier development of the form

$$G_s(\tau_1, \tau_2) = \frac{4\pi}{1-2s} E(\tau_2, s)\, v_1^{1-s} - 4\pi \sum_{n \neq 0} F_n(\tau_2, s)\, v_1^{\frac{1}{2}} K_{s-\frac{1}{2}}(2\pi|n|v_1)\, e^{-2\pi i n u_1},$$

where the series converges with exponential rapidity; here $E(\tau, s)$ is the Eisenstein series

$$E(\tau, s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \mathrm{Im}(\gamma\tau)^s = \frac{1}{2} \sum_{\substack{c,d \in \mathbb{Z} \\ (c,d)=1}} \frac{v^s}{|c\tau + d|^{2s}} \quad \left(\Gamma_\infty = \begin{pmatrix} 1 & \mathbb{Z} \\ 0 & 1 \end{pmatrix}\right),$$

$K_{s-\frac{1}{2}}$ is a K-Bessel function, and the $F_n(\tau_2, s)$ are meromorphic in s and holomorphic for $\mathrm{Re}(s) > \frac{1}{2}$. The Fourier expansion of $E(\tau, s)$ is

$$E(\tau, s) = v^s + \varphi(s)\, v^{1-s} + \frac{2\pi^s}{\Gamma(s)\,\zeta(2s)} \sum_{m \neq 0} |m|^{s-\frac{1}{2}} \sigma_{1-2s}(m)\, v^{\frac{1}{2}} K_{s-\frac{1}{2}}(2\pi|m|v)\, e^{2\pi i m u},$$

$$\varphi(s) = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(s-\frac{1}{2}\right)\zeta(2s-1)}{\Gamma(s)\,\zeta(2s)}, \quad \sigma_\nu(m) = \sum_{d|m} d^\nu.$$

Hence $G_s(\tau_1, \tau_2)$ can be meromorphically continued in s, the only pole in $\mathrm{Re}(s) > \frac{1}{2}$ being a simple one at $s=1$ with constant residue -12.

Using these properties, we can now prove

Proposition 5.1. *For τ_1, τ_2 two points of $\mathfrak{H}$ not equivalent under Γ we have the identity*

$$\log|j(\tau_1)-j(\tau_2)|^2 = \lim_{s\to 1}\left(G_s(\tau_1,\tau_2)+4\pi E(\tau_1,s)+4\pi E(\tau_2,s)-4\pi\varphi(s)\right)-24.$$

Proof. The limit exists by what was said above, since all four terms in the limit are meromorphic functions with simple poles at $s=1$, the residues being -12, 12, 12, and -12, respectively. We consider τ_2 as fixed and both sides of the asserted identity as functions of τ_1. Both are Γ-invariant. The function on the left is continuous in Γ except for a singularity $\log|\tau_1-\tau_2|^2+O(1)$ as $\tau_1 \to \tau_2$; by a), the function on the right has the same property. Both functions are harmonic; this is clear for the function on the left and follows for the function on the right by b), since

$$\lim_{s\to 1}\left(G_s(\tau_1,\tau_2)+4\pi E(\tau_1,s)\right)$$

is the limit of eigenfunctions of Δ_1 with eigenvalue $s(s-1)$ and

$$\lim_{s\to 1}\left(4\pi E(\tau_2,s)-4\pi\varphi(s)\right)-24$$

is constant (and hence harmonic) with respect to τ_1. Therefore it suffices to show that the two functions differ by $o(1)$ as $v_1 = \mathrm{Im}(\tau_1) \to \infty$. We have

$$\log|j(\tau_1)-j(\tau_2)|^2 = \log|e^{-2\pi i\tau_1}+O(1)|^2 = 4\pi v_1 + O(e^{-2\pi v_1}) \quad (v_1\to\infty),$$

while, by c) and the formula $K_{\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}\,e^{-x}$,

$$\begin{aligned}
&\lim_{s\to 1}\left(G_s(\tau_1,\tau_2)+4\pi E(\tau_1,s)+4\pi E(\tau_2,s)-4\pi\varphi(s)\right)\\
&\quad= 4\pi\lim_{s\to 1}\left(E(\tau_2,s)\left(1+\frac{1}{1-2s}v_1^{1-s}\right)\right)+4\pi\lim_{s\to 1}\left(E(\tau_1,s)-\varphi(s)\right)\\
&\qquad -2\pi\sum_{n\neq 0}|n|^{-\frac{1}{2}}F_n(\tau_2,1)\,e^{-2\pi|n|v_1}\,e^{-2\pi i n u_1}\\
&\quad= 12(\log v_1+2)+\left(4\pi v_1-12\log v_1+O(e^{-2\pi v_1})\right)+O(e^{-2\pi v_1}).
\end{aligned}$$

Thus the functions agree within $O(e^{-2\pi v_1})$ as $v_1\to\infty$ and the Proposition is proved.

We remark that the Proposition extends immediately to give a formula for the logarithm of the absolute value of the quantity $\varphi_m(j(\tau_1), j(\tau_2))$ defined in (4.2). Indeed, applying the m^{th} Hecke operator T_m with respect to τ_2 (i.e. replacing τ_2 by $\gamma\tau_2$,

where γ runs over a set of representatives modulo $SL_2(\mathbb{Z})$ of matrices of determinant m, and summing), and noting that $E(\tau,s)$ is an eigenfunction of T_m with eigenvalue $m^s\sigma_{1-2s}(m)$, we find

$$\begin{aligned}\log|\varphi_m(j(\tau_1),j(\tau_2))|^2 &= \lim_{s\to 1}\left(G_s^m(\tau_1,\tau_2)+4\pi\sigma_1(m)\,E(\tau_1,s)\right.\\ &\quad\left.+4\pi m^s\sigma_{1-2s}(m)\,E(\tau_1,s)-4\pi\sigma_1(m)\,\varphi(s)\right)-24\sigma_1(m)\\ &= \lim_{s\to 1}\left(G_s^m(\tau_1,\tau_2)+4\pi\sigma_1(m)\,E(\tau_1,s)+E(\tau_2,s)-\varphi(s)]\right)-24\sigma_1(m)\\ &\quad -12\sum_{d|m} d\log\frac{m}{d^2},\end{aligned}\tag{5.2}$$

where

$$G_s^m(\tau_1,\tau_2)=\frac{1}{2}\sum_{\substack{a,b,c,d\in\mathbb{Z}\\ ad-bc=m}} g_s\left(\tau_1,\frac{a\tau_2+b}{c\tau_2+d}\right).$$

We apply the Proposition to compute $\log|J(d_1,d_2)|^2$ where (as in Section 1) d_1 and d_2 are coprime negative fundamental discriminants and $J(d_1,d_2)$ is defined by (1.2). Let K_j be $\mathbb{Q}(\sqrt{d_j})$, $h_j=h(d_j)$ the class number of K_j, and $w_j(=2,\ 4$ or $6)$ the number units of K_j; then

$$\log|J(d_1,d_2)|^2=\frac{2}{w_1}\frac{2}{w_2}\sum_{[\tau_1],[\tau_2]}\log|j(\tau_1)-j(\tau_2)|^2$$

where the sum is over the h_1h_2 pairs of points $\tau_1,\tau_2\in\Gamma\backslash\mathfrak{H}$ $(\Gamma=PSL_2(\mathbb{Z}))$ of discriminant d_1,d_2. Let $\Gamma_{\tau_j}\subset\Gamma$ be the stabilizer of τ_j; then

$$\begin{aligned}\frac{2}{w_1}\frac{2}{w_2}\sum_{[\tau_1],[\tau_2]\in\Gamma\backslash\mathfrak{H}} G_s(\tau_1,\gamma\tau_2) &= \sum_{\substack{[\tau_1],[\tau_2]\in\Gamma\backslash\mathfrak{H}\\ \operatorname{disc}\tau_j=d_j}}\ \sum_{\gamma\in\Gamma_{\tau_1}\backslash\Gamma/\Gamma_{\tau_2}} g_s(\tau_1,\gamma\tau_2)\\ &= \sum_{\substack{[\tau_1],[\tau_2]\in\Gamma\backslash\mathfrak{H}\\ \operatorname{disc}\tau_j=d_j}}\ \sum_{(\gamma_1,\gamma_2)\in\Gamma\backslash(\Gamma\backslash\Gamma_{\tau_1}\times\Gamma/\Gamma_{\tau_2})} g_s(\gamma_1\tau_1,\gamma_2\tau_2),\end{aligned}$$

where we have written $\gamma=\gamma_1^{-1}\gamma_2$ with $\gamma_1,\gamma_2\in\Gamma$ well-defined up to right multiplication by elements of $\Gamma_{\tau_1},\Gamma_{\tau_2}$ and up to simultaneous left multiplication by an element of Γ. The set of $\gamma_j\tau_j$ as $[\tau_j]$ ranges over a set of representatives for Γ-equivalence classes of $\tau\in\mathfrak{H}$ with discriminant d_j and γ_j over $\Gamma_{\tau_j}\backslash\Gamma$ is simply the set of all points in $\mathfrak{H}$ with discriminant d_j. Hence

$$\frac{2}{w_1}\frac{2}{w_2}\sum_{[\tau_1],[\tau_2]} G_s(\tau_1,\tau_2)=\sum_{\substack{(\tau_1,\tau_2)\in\Gamma\backslash\mathfrak{H}^2\\ \operatorname{disc}\tau_j=d_j}} g_s(\tau_1,\tau_2).$$

The points $\tau_j\in\mathfrak{H}$ of discriminant d_j are in $1-1$ correspondence with the positive definite binary quadratic forms $Q_j(x,y)=a_jx^2+b_jxy+c_jy^2$ of discriminant $b_j^2-4a_jc_j$

the correspondence being that $\tau_j = \dfrac{-b_j+\sqrt{d_j}}{2a_j}$ is the root of $Q_j(\tau, 1)=0$ with positive imaginary part. Under this correspondence we have

$$g_s(\tau_1, \tau_2) = -2Q_{s-1}\left(\frac{n}{\sqrt{D}}\right) \text{ where } D = d_1 d_2 \text{ and } n = 2a_1 c_2 + 2a_2 c_1 - b_1 b_2.$$

Hence

$$\frac{2}{w_1}\frac{2}{w_2} G_s(\tau_1, \tau_2) = -2 \sum_{\substack{n > \sqrt{D} \\ n \equiv D (\mathrm{mod}\ 2)}} \rho(n)\, Q_{s-1}\left(\frac{n}{\sqrt{D}}\right)$$

where

$$\rho(n) = \frac{1}{2} \#\{(Q_1, Q_2) \in \mathfrak{Q}^2/\Gamma \mid \Delta(Q_j) = d_j,\ B_\Delta(Q_1, Q_2) = -n\};$$

here $\mathfrak{Q} \approx \mathbb{Z}^3$ is the set of all integral binary quadratic forms with the usual action of Γ, $\Delta : \mathfrak{Q} \to \mathbb{Z}$ the discriminant function, and B_Δ the associated bilinear form

$$B_\Delta([a_1, b_1, c_1], [a_2, b_2, c_2]) = b_1 b_2 - 2a_1 c_2 - 2a_2 c_1;$$

the factor $\frac{1}{2}$ arises because two forms Q_1, Q_2 satisfying the conditions given (with $d_1, d_2 < 0$, $n > 0$) are either both positive definite or both negative definite and we want only the first case. The three conditions $\Delta(Q_1) = d_1$, $\Delta(Q_2) = d_2$, $B_\Delta(Q_1, Q_2) = -n$ are equivalent to

$$\Delta(\xi Q_1 + \eta Q_2) = d_1 \xi^2 - 2n\xi\eta + d_2 \eta^2,$$

i.e. (Q_1, Q_2) yield a representation of the indefinite binary quadratic form $[d_1, -2n, d_2]$ by the ternary quadratic form Δ. Since the automorphism group of $(\mathfrak{Q}, \Delta)$ is $\{\pm 1\} \times \Gamma$, $\rho(n)$ is simply the number of inequivalent representations of $[d_1, -2n, d_2]$ by Δ. On the other hand, we have

$$E(\tau_j, s) = \frac{w_j}{2} \left|\frac{d_j}{4}\right|^{-\frac{s}{2}} \zeta(2s)^{-1}\, \zeta_{K_j, \mathscr{A}_j}(s),$$

where $\mathscr{A}_j$ is the ideal class of K_j corresponding to $[\tau_j] \in \Gamma\backslash\mathfrak{H}$ and $\zeta_{K_j, \mathscr{A}_j}$ the corresponding zeta-function (the sum of $\mathrm{N}(\mathfrak{a})^{-s}$ for all integral ideals $\mathfrak{a} \in \mathscr{A}_j$). The sum of the $\zeta_{K_j, \mathscr{A}_j}$ over all ideal classes $\mathscr{A}_j$ is the Dedekind zeta-function $\zeta_{K_j}(s)$. Hence we have proved:

Proposition 5.3. *Let K_1, K_2 be two imaginary quadratic fields with coprime discriminants d_1, d_2 and $J(d_1, d_2)$ the number defined by* (1.2). *Then*

$$\log |J(d_1, d_2)|^2 = \lim_{s \to 1} \left[-2 \sum_{\substack{n > \sqrt{D} \\ n \equiv D (\mathrm{mod}\ 2)}} \rho(n)\, Q_{s-1}\left(\frac{n}{\sqrt{D}}\right) + \frac{4\pi}{\zeta(2s)} \left(h_2' \left|\frac{d_1}{4}\right|^{\frac{s}{2}} \zeta_{K_1}(s) \right.\right.$$

$$\left.\left. + h_1' \left|\frac{d_2}{4}\right|^{\frac{s}{2}} \zeta_{K_2}(s) - h_1' h_2' \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(s - \frac{1}{2}\right)}{\Gamma(s)} \zeta(2s-1) \right) \right] - 24 h_1' h_2',$$

where $D = d_1 d_2$, $h_j' = \dfrac{2}{w_j} h_j \left(= \frac{1}{2} \text{ or } \frac{1}{3} \text{ if } d_j = -4 \text{ or } -3 \text{ and } h(K_j) \text{ otherwise} \right)$ and $\rho(n)$ is the number of inequivalent representations of the binary quadratic form $[d_1, -2n, d_2]$ by the form $\Delta = b^2 - 4ac$ on $\mathfrak{Q} = \{[a, b, c] \mid a, b, c \in \mathbb{Z}\}$.

Using (5. 2) instead of Proposition 5. 1 we can give a similar formula for

$$\sum_{[\tau_1],[\tau_2]} \log|\varphi_m(\tau_1, \tau_2)|^2$$

instead of $\log|J(d_1, d_2)|^2$, the argument of G_{s-1} now being $\dfrac{n}{m\sqrt{D}}$ for some $n > m\sqrt{D}$, $n \equiv mD \pmod 2$. The details are left to the reader.

6. The formula for $\log|J(d_1, d_2)|^2$ obtained in section 5 is not yet very useful because $\rho(n)$ is expressed as the number of orbits of an infinite set by an infinite group. In this section we will give an expression for $\rho(n)$ as a finite sum.

Proposition 6. 1. *Let d_1, d_2, D and $\rho(n)$ $(n > \sqrt{D}$, $n \equiv D \pmod 2)$ be as in the last proposition. Then*

$$\rho(n) = \sum_{d \mid \frac{n^2-D}{4}} \varepsilon(d),$$

where $\varepsilon(d)$ for integers $d > 0$ such that D is congruent to a square modulo $4d$ is defined as in Section 1 (namely as the multiplicative function which for primes l equals whichever of $\left(\dfrac{d_1}{l}\right)$ and $\left(\dfrac{d_2}{l}\right)$ is non-zero).

Proof. Let $K = \mathbb{Q}(\sqrt{D})$ be the real quadratic field of discriminant D, and χ the genus character of K corresponding to the decomposition $D = d_1 \cdot d_2$. We recall that χ is a character from the narrow class group to ± 1 with $\chi(\mathfrak{p}) = 1$ if $\mathfrak{p}$ is an inert prime ideal and $\chi(\mathfrak{p}) = \varepsilon(\mathrm{N}\mathfrak{p})$ otherwise; because χ corresponds to a decomposition of D into negative factors, we have $\chi(\mathfrak{a}) = -1$ if $\mathfrak{a}$ is a principal ideal generated by an element of negative norm. Let $\mu = \dfrac{n - \sqrt{D}}{2} \in \mathcal{O}_K$; then the principal ideal (μ) has norm $\dfrac{n^2 - D}{4}$ and is primitive (not divisible by a natural number > 1) because the coefficient of $\sqrt{D}$ in 2μ is 1, and one easily deduces that

$$\sum_{d \mid \frac{n^2-D}{4}} \varepsilon(d) = \sum_{\mathfrak{b} \mid (\mu)} \chi(\mathfrak{b}).$$

Let $L = \mathbb{Q}(\sqrt{d_1}, \sqrt{d_2})$; then L/K is the unramified quadratic extension corresponding to the character χ, so

$$\sum_{\mathfrak{b} \mid \mathfrak{a}} \chi(\mathfrak{b}) = r_{L/K}(\mathfrak{a})$$

for any integral ideal $\mathfrak{a}$ of K, where $r_{L/K}(\mathfrak{a})$ is the number of integral ideals $\mathfrak{A}$ of L with $\mathrm{N}_{L/K}(\mathfrak{A}) = \mathfrak{a}$. Therefore the identity to be proved is

$$\rho(n) = r_{L/K}((\mu)), \quad \mu = \frac{n - \sqrt{D}}{2}, \tag{6. 2}$$

i.e. we would like to establish a correspondence between the representations of $[d_1, -2n, d_2]$ by Δ and the ideals $\mathfrak{A}$ of L with norm (μ).

There is a one-to-one correspondence between positive definite binary quadratic forms Q_j of discriminant d_j and triples $(\mathfrak{a}_j, \alpha_j, \beta_j)$ modulo the action of K_j^*, where $\mathfrak{a}_j$ is a fractional ideal of K_j, (α_j, β_j) an oriented $\mathbb{Z}$-basis of $\mathfrak{a}_j$ (i.e. one with $\mathrm{Im}(\alpha_j\overline{\beta_j}) > 0$) and K_j^* acts by $\lambda(\mathfrak{a}_j, \alpha_j, \beta_j) = (\lambda\mathfrak{a}_j, \lambda\alpha_j, \lambda\beta_j)$; this correspondence associates to $(\mathfrak{a}_j, \alpha_j, \beta_j)$ the quadratic form $Q_j(x, y) = \dfrac{\mathrm{N}(\alpha_j x + \beta_j y)}{\mathrm{N}(\mathfrak{a}_j)}$. The action of $SL_2(\mathbb{Z})$ on quadratic forms corresponds to the action of $SL_2(\mathbb{Z})$ on oriented bases:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ (\mathfrak{a}_j, \alpha_j, \beta_j) = (\mathfrak{a}_j, a\alpha_j + b\beta_j, c\alpha_j + d\beta_j).$$

If $Q_1 = [a_1, b_1, c_1]$ and $Q_2 = [a_2, b_2, c_2]$ are forms of discriminant d_1 and d_2 corresponding to $(\mathfrak{a}_1, \alpha_1, \beta_1)$ and $(\mathfrak{a}_2, \alpha_2, \beta_2)$, and $B_4(Q_1, Q_2) = -n$, then (denoting conjugation in L/K or $K_j/\mathbb{Q}$ by $'$) we have:

$$\begin{aligned}
\mathrm{N}_{L/K}(\alpha_1\beta_2 - \alpha_2\beta_1) &= (\alpha_1\beta_2 - \alpha_2\beta_1)(\alpha_1'\beta_2' - \alpha_2'\beta_1') \\
&= \mathrm{N}(\alpha_1)\,\mathrm{N}(\beta_2) + \mathrm{N}(\alpha_2)\,\mathrm{N}(\beta_1) - \frac{1}{2}(\alpha_1\beta_1' + \alpha_1'\beta_1)(\alpha_2\beta_2' + \alpha_2'\beta_2) \\
&\quad + \frac{1}{2}(\alpha_1\beta_1' - \alpha_1'\beta_1)(\alpha_2\beta_2' - \alpha_2'\beta_2) \\
&= \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\left(a_1c_2 + a_2c_1 - \frac{1}{2}b_1b_2 - \frac{1}{2}\sqrt{d_1d_2}\right) \\
&= \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\,\frac{n - \sqrt{D}}{2}.
\end{aligned}$$

Hence

$$\begin{aligned}
\rho(n) &= \#\{(Q_1, Q_2) \in \mathfrak{Q}^2/\Gamma \mid Q_j \text{ positive definite, } \mathrm{disc}\, Q_j = d_j,\ B_4(Q_1, Q_2) = -n\} \\
&= \#\{((\mathfrak{a}_1, \alpha_1, \beta_1), (\mathfrak{a}_2, \alpha_2, \beta_2)) \bmod K_1^* \times K_2^* \times SL_2(\mathbb{Z}) \mid \mathrm{N}_{L/K}(\alpha_1\beta_2 - \alpha_2\beta_1) \\
&\qquad = \mu\mathrm{N}(\alpha_1)\,\mathrm{N}(\alpha_2)\},
\end{aligned}$$

where $\mathfrak{a}_j, \alpha_j, \beta_j$ $(j = 1, 2)$ are as above, $SL_2(\mathbb{Z})$ acts simultaneously on (α_1, β_2) and (α_2, β_2), and $\mu = \dfrac{n - \sqrt{D}}{2} \in \mathcal{O}_K$. Write $\mathfrak{a}_1\mathfrak{a}_2$ for the set of $\mathbb{Z}$-linear combinations of elements v_1v_2 $(v_j \in \mathfrak{a}_j)$ and ρ for the element $\alpha_1\beta_2 - \alpha_2\beta_1$ of L. Then $\rho \in \mathfrak{a}_1\mathfrak{a}_2$ and

$$\mathrm{N}_{L/K}(\rho) = \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\,\mu.$$

Conversely, any element ρ of $\mathfrak{a}_1\mathfrak{a}_2$ with $\mathrm{N}(\rho) = \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\,\mu$ has the form $\alpha_1\beta_2 - \alpha_2\beta_1$ for some oriented bases (α_1, β_1) and (α_2, β_2) of $\mathfrak{a}_1$ and $\mathfrak{a}_2$. Indeed, choose an arbitrary oriented basis (α_1, β_1) of $\mathfrak{a}_1$; then $\rho \in \mathfrak{a}_1\mathfrak{a}_2$ implies $\rho = \alpha_1\beta_2 - \alpha_2\beta_1$ with some $\alpha_2, \beta_2 \in \mathfrak{a}_2$ and the fact that the coefficient of $\sqrt{D}$ in $\dfrac{\mathrm{N}(\rho)}{\mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)}$ is $-\dfrac{1}{2}$ implies that (α_2, β_2) is an oriented basis of $\mathfrak{a}_2$ $\left(\text{it would be } \pm\dfrac{N}{2} \text{ if } \mathbb{Z}\alpha_2 + \mathbb{Z}\beta_2 \text{ had index } N \text{ in } \mathfrak{a}_2 \text{ and } +\dfrac{1}{2} \text{ if the basis were unoriented}\right)$. The same argument shows that ρ determines (α_2, β_2) uniquely

given (α_1, β_1); since all oriented bases of $\mathfrak{a}_1$ differ by elements of $SL_2(\mathbb{Z})$, the choice of ρ uniquely determines *both* oriented bases (α_j, β_j) up to the simultaneous action of $SL_2(\mathbb{Z})$. Hence

$$\rho(n) = \#\{(\mathfrak{a}_1, \mathfrak{a}_2, \rho) \mid \mathfrak{a}_j \text{ a fractional ideal of } K_j,\ \rho \in \mathfrak{a}_1, \mathfrak{a}_2,$$
$$\mathrm{N}_{L/K}(\rho) = \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\,\mu\}/K_1^* \times K_2^*,$$

where $K_1^* \times K_2^*$ acts by $(\mathfrak{a}_1, \mathfrak{a}_2, \rho) \to (\lambda_1 \mathfrak{a}_1, \lambda_2 \mathfrak{a}_2, \lambda_1 \lambda_2 \rho)$ $(\lambda_j \in K_j^*)$. The freedom of choosing λ_1 and λ_2 means that we can fix the choice of $\mathfrak{a}_1$ and $\mathfrak{a}_2$ within their ideal classes; then we still have the freedom of changing λ_j by a unit of K_j. Hence

$$\rho(n) = \sum_{\substack{[\mathfrak{a}_1] \in C_{K_1} \\ [\mathfrak{a}_2] \in C_{K_2}}} \#\{\rho \in \mathfrak{a}_1 \mathfrak{a}_2 / U_{K_1} \cdot U_{K_2} \mid \mathrm{N}_{L/K}(\rho) = \mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)\,\mu\}, \tag{6.3}$$

where C_{K_j} and U_{K_j} denote the class and unit groups of $K_j (j = 1,2)$, $\mathfrak{a}_j (j = 1,2)$ is any (fractional) ideal of K_j in the class $[\mathfrak{a}_j]$, and $\mathfrak{a}_1 \mathfrak{a}_2$ denotes the set of $\mathbb{Z}$-linear combinations of elements $\theta_1 \theta_2$ with $\theta_j \in \mathfrak{a}_j$ (it is clear that the summand depends only on the classes of $\mathfrak{a}_1$ and $\mathfrak{a}_2$). Let C_L and U_L denote the class and unit groups of the biquadratic field L and C_K^+ and U_K^+ the strict ideal class group and group of totally positive units of K. Then we have the exact sequence

$$0 \to \{\pm 1\} \to U_{K_1} \times U_{K_2} \to U_L \xrightarrow{\mathrm{N}} U_K^+ \to C_{K_1} \times C_{K_2} \to C_L \xrightarrow{\mathrm{N}} C_K^+ \xrightarrow{\chi} \{\pm 1\} \to 0.$$

This is proved by a standard argument using elementary class field theory and the analytic class number formulas for K_1, K_2, K and L as in Hasse [11] (esp. § 26); we omit the proof and the definition of the map $U_K^+ \to C_{K_1} \times C_{K_2}$, which depends on genera theory. Using the exact sequence, we can establish a 1:1 correspondence between the triples $([\mathfrak{a}_1], [\mathfrak{a}_2], \rho)$ counted in (6.3) and the integral ideals $\mathfrak{A}$ of L with norm (μ), establishing (6.2). Indeed, because d_1 and d_2 are coprime we have $\mathcal{O}_L \cong \mathcal{O}_{K_1} \otimes_{\mathbb{Z}} \mathcal{O}_{K_2}$, so $\mathfrak{a}_1 \mathfrak{a}_2$ is a (fractional) ideal of L and $\mathfrak{A} = \rho^{-1} \mathfrak{a}_1 \mathfrak{a}_2$ for ρ as in (6.3) is an integral ideal with $\mathrm{N}_{L/K}(\mathfrak{A}) = (\mu)$. Conversely, let $\mathfrak{A}$ be an integral ideal with norm (μ). Since μ has positive norm, the ideal class $[\mathfrak{A}]$ is in the kernel of $\mathrm{N}_{L/K}\colon C_L \to C_K^+$, so the exactness of the sequence at C_L implies the existence of ideals $\mathfrak{a}_1$, $\mathfrak{a}_2$ with $\mathfrak{A} \sim \mathfrak{a}_1 \mathfrak{a}_2$, and the exactness at $C_{K_1} \times C_{K_2}$ and U_K^+ implies that there are exactly $Q = [U_K^+ \colon \mathrm{N}_{L/K}(U_L)]$ $(= 1$ or $2)$ choices for $([\mathfrak{a}_1], [\mathfrak{a}_2])$. From $\mathfrak{A} \sim \mathfrak{a}_1 \mathfrak{a}_2$ we have $\mathfrak{A}^{-1} \mathfrak{a}_1 \mathfrak{a}_2 = (\rho)$ for some $\rho \in L$, and then $\mathrm{N}(\mathfrak{A}) = (\mu)$ implies that $\dfrac{\mathrm{N}_{L/K}(\rho)}{\mathrm{N}\mathfrak{a}_1 \mathrm{N}\mathfrak{a}_2} = \varepsilon\mu$ for some unit ε of K. Since $\mathrm{N}_{L/K}(\rho)$ and μ are totally positive, $\varepsilon \in U_K^+$. Among the Q choices of $([\mathfrak{a}_1], [\mathfrak{a}_2])$, exactly one will correspond to $\varepsilon \in \mathrm{N}_{L/K}(U_L)$. We make this choice; then ρ can be modified by a unit of U_L to achieve $\varepsilon = 1$, i.e. $\dfrac{\mathrm{N}_{L/K}(\rho)}{\mathrm{N}(\mathfrak{a}_1)\,\mathrm{N}(\mathfrak{a}_2)} = \mu$, the choice of ρ now being unique up to an element of $U_{K_1} \cdot U_{K_2}$ (exactness at U_L). This completes the proof.

7. The result we want to prove, Theorem 1.3, can be written

$$-\log |J(d_1, d_2)|^2 = \sum_{\substack{x^2 < D \\ x^2 \equiv D \pmod 4}} \sum_{n \mid \frac{x^2 - D}{4}} \varepsilon(n) \log n$$
$$= \sum_{\substack{\nu \in \mathfrak{b}^{-1} \\ \nu \gg 0 \\ \mathrm{Tr}(\nu) = 1}} \sum_{n \mid (\nu)\mathfrak{b}} \chi(n) \log \mathrm{N}(n), \tag{7.1}$$

where again we have written D for $d_1 \cdot d_2$ and χ for the corresponding genus character on $K = \mathbb{Q}(\sqrt{D})$ and $\mathfrak{d} = (\sqrt{D})$ is the different of K; the second line follows from the first on setting $\nu = \dfrac{x+\sqrt{D}}{2\sqrt{D}}$ and noticing that the correspondence $\mathfrak{n} \to \mathrm{N}(\mathfrak{n})$ gives a bijection between the ideal divisors of the primitive integral ideal $(\nu\sqrt{D})$ and the positive divisors of $\dfrac{D-x^2}{4}$, with $\chi(\mathfrak{n}) = \varepsilon(n)$. Formula (7.1) is very reminiscent of the formulas

$$30k\,\zeta_K(-k+1) = \sum_{\substack{\nu \in \mathfrak{d}^{-1} \\ \nu \gg 0 \\ \mathrm{Tr}(\nu)=1}} \sum_{\mathfrak{n} \mid (\nu)\mathfrak{d}} \mathrm{N}(\mathfrak{n})^{k-1} \quad (k = 2, 4)$$

of Siegel ([18], see also [21] or [4]), the only difference being that $\mathrm{N}(\mathfrak{n})^{k-1}$ is replaced by $\chi(\mathfrak{n}) \log \mathrm{N}(\mathfrak{n})$. Siegel's formulas came from restricting to the diagonal $z = z'$ the Hecke-Eisenstein series

$$E_{K,k}(z, z') = \sum_{[\mathfrak{a}] \in C_K} \mathrm{N}(\mathfrak{a})^k \sum_{\substack{(m,n) \in \mathfrak{a}^2/\mathcal{O}^\times \\ (m,n) \neq (0,0)}} \frac{1}{(mz+n)^k\,(m'z'+n')^k} \qquad (z, z' \in \mathfrak{H})$$

of weight k on $SL_2(\mathcal{O}_K)$ and identifying the resulting modular form of weight $2k$ on $SL_2(\mathbb{Z})$ with a multiple of $E_{2k}(z)$. Thus the term $\mathrm{N}(\mathfrak{n})^{k-1}$ corresponds to a holomorphic Eisenstein series of weight k on $SL_2(\mathcal{O}_K)$, so one can except the analogous formula with $\chi(\mathfrak{n}) \log \mathrm{N}(\mathfrak{n})$ to be related to the function $\dfrac{\partial}{\partial s} E_s(z, z)|_{s=0}$, where

$$E_s(z, z') = E_{K,\chi,1,s}(z, z')$$

$$(7.2) \quad = \sum_{[\mathfrak{a}] \in C_K} \chi(\mathfrak{a})\,\mathrm{N}(\mathfrak{a})^{1+2s} \sum_{\substack{(m,n) \in \mathfrak{a}^2/\mathcal{O}^\times \\ (m,n) \neq (0,0)}} \frac{y^2 y'^s}{(mz+n)\,(m'z'+n')|mz+n|^{2s}|m'z'+n'|^{2s}}$$

$$(z = x+iy,\ z' = x'+iy' \in \mathfrak{H})$$

is the non-holomorphic Eisenstein series of weight 1 on $SL_2(\mathcal{O}_K)$ introduced by Hecke [12]. Hecke's purpose was to produce a holomorphic Eisenstein series of weight 1 by introducing the factor $\dfrac{y^s y'^s}{|mz+n|^{2s}|m'z'+n'|^{2s}}$ $(\mathrm{Re}(s) > 0)$ into the non-convergent Eisenstein series and then letting $s \to 0$ ("Hecke's trick"). By computing the constant term of the limit, he thought he had shown that the function obtained at $s = 0$ was different from zero, but as is well-known (cf. Schoeneberg's corrections on p. 949 of Hecke's "Werke") the computation is invalidated by an error of sign and in fact the functions obtained by letting $s \to 0$ always vanish identically. This fact, unfortunate for Hecke, is very fortunate for us, for it means that the derivative $\dfrac{\partial}{\partial s} E_s|_{s=0}$ is the leading term of the Taylor expansion of E_s at $s = 0$ and therefore computable. We now describe this.

We begin by noting that C_K in (7.2) can be taken to be the wide ideal class group of K, because replacing $\mathfrak{a}$ by $\lambda\mathfrak{a}$ $(\lambda \in K^*)$ changes the inner sum by a factor

$$\frac{1}{\mathrm{N}(\lambda)|\mathrm{N}(\lambda)|^{2s}} = \mathrm{sgn}\,(\mathrm{N}(\lambda))|\mathrm{N}(\lambda)|^{-1-2s},$$

while $\chi(\mathfrak{a})$ changes by a factor $\operatorname{sgn}(\mathsf{N}(\lambda))$ (because χ is a genus character corresponding to a decomposition of D into *negative* factors) and $\mathsf{N}(\mathfrak{a})^{1+2s}$ by $|\mathsf{N}(\lambda)|^{1+2s}$. If K had a unit of negative norm, the series would vanish identically, but this cannot be the case for $D=d_1\cdot d_2$. Following Hecke—the method is by now quite standard—we find the Fourier expansion of $E_s(z,z')$:

$$E_s(z,z') = L_K(1+2s,\chi)\, y^s y'^s + D^{-\frac{1}{2}} L_K(s,\chi)\, \Phi_s(0)^2\, y^{-s} y'^{-s}$$
$$+ D^{-\frac{1}{2}} y^{-s} y'^{-s} \sum_{\substack{\nu\in\mathfrak{d}^{-1}\\ \nu\neq 0}} \sigma_{-2s,\chi}((\nu)\mathfrak{d})\, \Phi_s(\nu y)\, \Phi_s(\nu' y')\, e^{2\pi i(\nu x+\nu' x')},$$

where $L_K(s,\chi)=L\left(s,\left(\frac{d_1}{\cdot}\right)\right)L\left(s,\left(\frac{d_2}{\cdot}\right)\right)$ is the L-series of χ and

$$\Phi_s(t)=\int_{-\infty}^{\infty}\frac{e^{-2\pi i x t}}{(x+i)\,(x^2+1)^s}\,dx\ \ (t\in\mathbb{R}),\quad \sigma_{s,\chi}(\mathfrak{a})=\sum_{\mathfrak{n}|\mathfrak{a}}\chi(\mathfrak{n})\,\mathsf{N}(\mathfrak{n})^s.$$

By deforming the path of integration we see that $\Phi_s(t)$ has an analytic continuation to all s and is bounded (uniformly for s in compact sets) by $|t|^{O(1)}e^{-2\pi|t|}$ as $|t|\to\infty$, so this gives the holomorphic continuation of $E_s(z,z')$ to all $s\in\mathbb{C}$. At $s=0$ we have

$$\Phi_0(t)=\begin{cases}-2\pi i\, e^{-2\pi t} & t>0,\\ -\pi i & t=0,\\ 0 & t<0,\end{cases}$$

so the coefficients of $E_0(z,z')$ with ν not totally positive vanish. On the other hand, the constant term $L_K(1,\chi)-\pi^2 D^{-\frac{1}{2}} L_K(0,\chi)$ of $E_0(z,z')$ vanishes by the functional equation of $L_K(s,\chi)$ (this is the fact that Hecke's mistake of sign caused him to miss), while the terms with $\nu\gg 0$ vanish because the contributions of $\mathfrak{n}$ and $(\nu)\mathfrak{d}^{-1}\mathfrak{n}$ cancel (this was also overlocked by Hecke; cf. his remarks on pp. 386 and 394 of [12]). This shows that $E_s(z,z')$ vanishes at $s=0$ and also permits us to calculate its derivative there:

$$\frac{\partial}{\partial s}E_s(z,z')|_{s=0} = 2L_K(1,\chi)\log(yy')+4C_\chi$$
$$+8\pi^2 D^{-\frac{1}{2}}\sum_{\substack{\nu\in\mathfrak{d}^{-1}\\ \nu\gg 0}}\sigma'_\chi((\nu)\mathfrak{d})\,e^{2\pi i(\nu z+\nu' z')}$$
$$-4\pi^2 D^{-\frac{1}{2}}\sum_{\substack{\nu\in\mathfrak{d}^{-1}\\ \nu>0>\nu'}}\sigma_{0,\chi}((\nu)\mathfrak{d})\,\Phi(|\nu'|y')\,e^{2\pi i(\nu z+\nu' z')}$$
$$-4\pi^2 D^{-\frac{1}{2}}\sum_{\substack{\nu\in\mathfrak{d}^{-1}\\ \nu<0<\nu'}}\sigma_{0,\chi}((\nu)\mathfrak{d})\,\Phi(|\nu|y)\,e^{2\pi i(\nu z+\nu' z')}$$

with

$$C_\chi = L'_K(1,\chi)+\left(\frac{1}{2}\log D-\log\pi-\gamma\right)L_K(1,\chi)\ \ (\gamma=\text{Euler's constant}),$$

$$\sigma'_\chi(\mathfrak{a})=\frac{\partial}{\partial s}\sigma_{s,\chi}(\mathfrak{a})|_{s=0}=\sum_{\mathfrak{n}|\mathfrak{a}}\chi(\mathfrak{n})\log\mathsf{N}(\mathfrak{n}),$$

$$\Phi(t)=\frac{i}{2\pi}\,e^{-2\pi t}\,\frac{\partial}{\partial s}\Phi_s(-t)\Big|_{s=0}.$$

(The terms with $v \ll 0$ contribute nothing because $\Phi_s(vy)$, $\Phi_s(v'y')$ and $\sigma_{s,\chi}((v)\mathfrak{d})$ all vanish at $s=0$, so the corresponding Fourier coefficients of E_s have a third-order zero.) Therefore the function

$$F(z) = \frac{\sqrt{D}}{8\pi^2} \frac{\partial}{\partial s} E_s(z, z)\Big|_{s=0} \qquad (z \in \mathfrak{H})$$

has the Fourier expansion

$$F(z) = \frac{\sqrt{D}}{2\pi^2} (L_K(1, \chi) \log y + C_\chi) + \sum_{\substack{v \in \mathfrak{d}^{-1} \\ v \gg 0}} \sigma'_\chi((v)\mathfrak{d})\, e^{2\pi i \operatorname{Tr}(v) z}$$
$$- \sum_{\substack{v \in \mathfrak{d}^{-1} \\ v > 0 > v'}} \sigma_{0,\chi}((v)\mathfrak{d})\, \Phi(|v'|y)\, e^{2\pi i \operatorname{Tr}(v) z}$$

(convergent because $\Phi(t) = O(e^{-4\pi t})$ as $t \to \infty$). Now we apply to this the following result.

Proposition 7.3. *Let $F(z)$ be a function on $\mathfrak{H}$ which transforms under $SL_2(\mathbb{Z})$ like a modular form of weight 2 and satisfies $F(z) = A \log y + B + O(y^{-\varepsilon})$ as $y \to \infty$ for some constants A, B and $\varepsilon > 0$. Let the Fourier expansion of $F(z)$ be $\sum_{m=-\infty}^{\infty} a_m(y)\, e^{2\pi i m z}$. Then*

$$\lim_{s \to 0} \left(4\pi \int_0^\infty a_1(y)\, e^{-4\pi y}\, y^s\, dy + \frac{24A}{s} \right) = 24A \left(2\frac{\zeta'}{\zeta}(2) + 1 + \log 4 \right) - 24B.$$

For $m > 1$ there is an analogous formula for

$$\lim_{s \to 0} \left(4\pi m \int_0^\infty a_m(y)\, e^{-4\pi m y}\, y^s\, dy + \frac{24A\sigma(m)}{s} \right),$$

where $\sigma(m)$ denotes the sum of the divisors of m.

This result is an extension of a result of Sturm on holomorphic projections of modular forms [19]. We do not give the details of the proof, since a more general result (for forms of arbitrary weight and level) is given in [10], but merely sketch the idea. For $\operatorname{Re}(s) > 0$ the m^{th} non-holomorphic Poincaré series of weight 2 is defined by

$$P_{2,s}^{(m)}(z) = \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_\infty \backslash \Gamma} \frac{e^{2\pi i m \frac{az+b}{cz+d}}}{(cz+d)^2} \frac{y^s}{|cz+d|^{2s}}.$$

If $F(z) = O(y^{-\varepsilon})$, then the Petersson scalar product of F and $P_{2,s}^{(m)}$ converges absolutely (even if the terms in the Poincaré series are replaced by their absolute values) and equals $\int_0^\infty a_m(y)\, e^{-4\pi m y}\, y^s\, dy$ for $\operatorname{Re}(s) > 0$. On the other hand, it is known that $P_{2,s}^{(m)}$ has an analytic continuation to $s=0$ and vanishes there (because there are no holomorphic modular forms of weight 2 on $SL_2(\mathbb{Z})$). This proves the proposition in the case $A=0$, $B=0$. For the general case it then suffices to consider a single function with $A=0$, $B \neq 0$ and one with $A \neq 0$, and taking the value and derivative at $s=0$ of the non-holomorphic Eisenstein series $E_{2,s} = P_{2,s}^{(0)}$ we obtain the formula given.

In our situation we have $A=\frac{\sqrt{D}}{2\pi^2}L_K(1,\chi)$, $B=\frac{\sqrt{D}}{2\pi^2}C_\chi$ and

$$a_1(y)=\sum_{\substack{v\in\mathfrak{d}^{-1}\\ v\gg 0\\ \mathrm{Tr}(v)=1}}\sigma'_\chi((v)\mathfrak{d})-\sum_{\substack{v\in\mathfrak{d}^{-1}\\ v>0>v'\\ \mathrm{Tr}(v)=1}}\sigma_{0,\chi}((v)\mathfrak{d})\,\Phi(|v'|y).$$

The first term, which we denote by S, is independent of y and equals the expression occurring on the right-hand side of (7. 1). In the second term we write $v=\frac{n+\sqrt{D}}{2\sqrt{D}}$; then $n>\sqrt{D}$, $n\equiv D\,(\mathrm{mod}\,2)$ and

$$\sigma_{0,\chi}((v)\mathfrak{d})=\sigma_{0,\chi}\left(\left(\frac{n+\sqrt{D}}{2}\right)\right)=\sum_{d\mid\frac{n^2-D}{4}}\varepsilon(d)=\rho(n)$$

by Proposition 6. 1. Therefore

$$4\pi\int_0^\infty a_1(y)\,e^{-4\pi y}y^s\,dy=\frac{\Gamma(s+1)}{(4\pi)^s}S-\sum_{\substack{n>\sqrt{D}\\ n\equiv D(2)}}\rho(n)\,\Psi_s\left(\frac{n-\sqrt{D}}{2\sqrt{D}}\right)$$

with

$$\Psi_s(\lambda)=4\pi\int_0^\infty\Phi(\lambda y)\,e^{-4\pi y}y^s\,dy\quad(\lambda>0),\tag{7.4}$$

and Proposition 7. 3 gives

$$\begin{aligned}S=\lim_{s\to 0}\left(\sum_{\substack{n>\sqrt{D}\\ n\equiv D(2)}}\rho(n)\,\Psi_s\left(\frac{n-\sqrt{D}}{2\sqrt{D}}\right)-\frac{12\sqrt{D}}{\pi^2}L_K(1,\chi)\,s^{-1}\right)\\ +\frac{12\sqrt{D}}{\pi^2}L_K(1,\chi)\left(2\frac{\zeta'}{\zeta}(2)+1+\log 4\right)-\frac{12\sqrt{D}}{\pi^2}C_\chi.\end{aligned}\tag{7.5}$$

To complete the proof we must calculate the function $\Psi_s(\lambda)$ defined by (7. 4) near $s=0$. First we need a formula for $\Phi(y)$. For $t>0$ we can deform the path of integration in the integral defining $\Phi_s(t)$ to a path C circling the positive imaginary axis from $-\varepsilon+i\infty$ to $+\varepsilon+i\infty$ in a counter-clockwise direction. The resulting integral is convergent for all $s\in\mathbb{C}$, so we obtain the holomorphic continuation (in s) of $\Phi_s(t)$ and—differentiating under the integral sign and setting $s=0$—the formula

$$\Phi(t)=\frac{-i}{2\pi}e^{-2\pi t}\int_C\frac{1}{x+i}\log(x^2+1)\,e^{2\pi ixt}\,dx\quad(t>0).$$

Since $\log(x^2+1)$ changes by $2\pi i$ as one crosses from the left to the right side of C, this is equal to

$$e^{-2\pi t}\int_i^{i\infty}\frac{1}{x+i}\,e^{2\pi ixt}\,dx.$$

Setting $x = i(2u-1)$, we obtain the formula

$$\Phi(t) = \int_1^\infty e^{-4\pi t u} \frac{du}{u} \quad (t > 0)$$

(exponential integral). Substituting this into (7.4) gives

$$\Psi_s(\lambda) = 4\pi \int_0^\infty \int_1^\infty e^{-4\pi u\lambda y} \frac{du}{u} e^{-4\pi y} y^s \, dy = \frac{\Gamma(s+1)}{(4\pi)^s} \int_1^\infty \frac{du}{u(1+\lambda\mu)^{s+1}}$$

for any $s \in \mathbb{C}$, $\mathrm{Re}(s) > -1$. In particular,

$$\Psi_0(\lambda) = \int_1^\infty \left(\frac{1}{u} - \frac{\lambda}{1+\lambda u}\right) du = \log \frac{u}{1+\lambda u}\bigg|_1^\infty = \log\left(1+\frac{1}{\lambda}\right)$$

and

$$\begin{aligned}\Psi_s(\lambda) &= \frac{\Gamma(s+1)}{(4\pi)^s} \int_1^\infty [\lambda^{-s-1} u^{-s-2} + O(\lambda^{-s-2} u^{-s-3})] \, du \\ &= \frac{1}{s+1} \frac{\Gamma(s+1)}{(4\pi)^s} \lambda^{-s-1} + O(\lambda^{-s-2}) \quad (\lambda \to \infty).\end{aligned}$$

On the other hand, from the definition of Q_{s-1} in terms of the hypergeometric definition given in § 5 we find

$$Q_0\left(\frac{1+t}{1-t}\right) = \frac{1}{2} \sum_{n=0}^\infty \frac{(1-t)^{n+1}}{n+1} = \frac{1}{2} \log \frac{1}{t} \quad (0 < t < 1),$$

$$Q_{s-1}\left(\frac{1+t}{1-t}\right) = \frac{\Gamma(s)^2}{2\Gamma(2s)} [(1-t)^s + O(1-t)^{s+1}] \quad (t \to 1)$$

or

$$Q_0(1+2\lambda) = \frac{1}{2} \log\left(1+\frac{1}{\lambda}\right),$$

$$Q_{s-1}(1+2\lambda) = \frac{\Gamma(s)^2}{2\Gamma(2s)} [\lambda^{-s} + O(\lambda^{-s-1})] \quad (\lambda \to \infty).$$

It follows that the function $\Psi_s(\lambda) - \dfrac{2\Gamma(2s+2)}{(4\pi)^s \Gamma(s+2)} Q_s(1+2\lambda)$ is $O(\lambda^{-s-2})$ as $\lambda \to \infty$ and vanishes identically for $s = 0$, so

$$\begin{aligned}&\lim_{s\to 0} \left[\sum_{\substack{n > \sqrt{D} \\ n \equiv D(2)}} \rho(n) \Psi_s\left(\frac{n-\sqrt{D}}{2\sqrt{D}}\right) - \frac{\kappa}{s} \right] \\ &= \lim_{s\to 1} \left[\frac{2\Gamma(2s)}{(4\pi)^{s-1} \Gamma(s+1)} \sum_{\substack{n > \sqrt{D} \\ n \equiv D(2)}} \rho(n) Q_{s-1}\left(\frac{n}{\sqrt{D}}\right) - \frac{\kappa}{s-1} \right] \\ &= \lim_{s\to 1} \left[2 \sum_{\substack{n > \sqrt{D} \\ n \equiv D(2)}} \rho(n) Q_{s-1}\left(\frac{n}{\sqrt{D}}\right) - \frac{(4\pi)^{s-1} \Gamma(s+1)}{\Gamma(2s)} \frac{\kappa}{s-1} \right].\end{aligned}$$

where κ is chosen to make the limit exist. Comparing this with (7. 5), we see that $\kappa = \frac{12\sqrt{D}}{\pi^2} L_K(1, \chi)$ and

$$S = \lim_{s \to 1} \left[\sum_{\substack{n > \sqrt{D} \\ n \not\equiv D(2)}} 2\rho(n) Q_{s-1}\left(\frac{n}{\sqrt{D}}\right) - \frac{\sqrt{D}}{\pi^2} \frac{L_K(1, \chi)}{s-1} \right]$$
$$+ \frac{12\sqrt{D}}{\pi^2} L_K(1, \chi) \left[2 + 2\frac{\zeta'}{\zeta}(2) - \frac{1}{2} \log D \right] - \frac{12\sqrt{D}}{\pi^2} L_K'(1, \chi).$$

Comparing this with Proposition 5. 3, and using the Taylor expansions

$$\zeta_{K_j}(s) = \zeta(s)\, L(s, \chi_j)$$
$$= \left[\frac{1}{s-1} + \gamma + \cdots \right] [L(1, \chi_j) + L'(1, \chi_j)(s-1) + \cdots] \quad (j = 1, 2),$$
$$L_K(s, \chi) = L(s, \chi_1)\, L(s, \chi_2)$$
$$= L(1, \chi_1)\, L(1, \chi_2) \left[1 + \left(\frac{L'}{L}(1, \chi_1) + \frac{L'}{L}(1, \chi_2) \right)(s-1) + \cdots \right],$$
$$L(1, \chi_j) = \frac{\pi}{\sqrt{|d_j|}} h_j' \quad (j = 1, 2),$$

we find

$$S = -\log |J(d_1, d_2)|^2$$

as was to be shown. This completes the analytic proof of the formula for $J(d_1, d_2)$. A similar calculation for the m^{th} coefficient of $F(z)$, using the general formula referred to in Proposition 7. 3 and the generalization of Proposition 5. 3 mentioned at the end of § 5, leads to a formula for $\sum_{[\tau_1], [\tau_2]} \log |\varphi_m(\tau_1, \tau_2)|^2$; we omit the details.

References

[1] *H. Bateman*, Higher Transcendental Functions. I, New York 1953.

[2] *W. E. H. Berwick*, Modular invariants, Proc. Lond. Math. Soc. **28** (1927), 53—69.

[3] *A. Borel, S. Chowla, C. S. Herz, K. Iwasawa* and *J-P. Serre*, Seminar on Complex Multiplication, Lecture Notes in Math. **21**, Berlin-Heidelberg-New York 1966.

[4] *H. Cohen*, Variations sur un thème de Siegel-Hecke, Bordeaux 1973/4, no. 5, 1—45.

[5] *M. Deuring*, Die Typen der Multiplikatorenringe elliptischer Funktionenkörper, Abh. Math. Sem. Hamburg **14** (1941), 197—272.

[6] *M. Deuring*, Teilbarkeitseigenschaften der singulären Moduln der elliptischen Funktionen und die Diskriminante der Klassengleichung, Comm. Math. Helv. **19** (1946), 74—82.

[7] *D. Dorman*, Prime factorization of singular moduli, Thesis, Brown 1984.

[8] *B. Gross*, Arithmetic on elliptic curves with complex multiplication, Lecture Notes in Math. **776**, Berlin-Heidelberg-New York 1980.

[9] *B. Gross, W. Kohnen* and *D. Zagier*, Heegner points and derivatives of L-series. II, in preparation.

[10] *B. Gross* and *D. Zagier*, Heegner points and derivatives of L-series, in preparation.

[11] *H. Hasse*, Über die Klassenzahl abelscher Zahlkörper, Berlin 1952.
[12] *E. Hecke*, Analytische Funktionen und algebraische Zahlen, zweiter Teil, Abh. Math. Sem. Hamburg **3** (1924), 231—236, Mathematische Werke, Göttingen 1970, 381—404.
[13] *D. Hejhal*, The Selberg Trace Formula for $PSL(2, R)$. II, Lecture Notes Math. **1001**, Berlin-Heidelberg-New York 1984.
[14] *O. Herrmann*, Über die Berechnung der Fourierkoeffizienten der Funktion $j(\tau)$, J. reine angew. Math. **274** (1973), 187—195.
[15] *C. Jensen* and *N. Yui*, Polynomials with D_p as Galois group, J. Number Th. **15** (1982), 347—375.
[16] *J-P. Serre*, Groupes p-divisibles (d'après J. Tate), Séminaire Bourbaki **318**, 1966/67, 1—14.
[17] *J-P. Serre* and *J. Tate*, Good reduction of abelian varieties, Ann. of Math. **88** (1968), 492—517.
[18] *C. L. Siegel*, Berechnung von Zetafunktionen an ganzzahligen Stellen. Nachr. Akad. Wiss. Göttingen, Math.-Phys. Klasse **2** (1968), 7—38.
[19] *J. Sturm*, Projections of C^∞ automorphic forms, Bull. Amer. Math. Soc. **2** (1980), 435—439.
[20] *J. Tate*, Arithmetic of elliptic curves, Inv. math. **23** (1974), 179—206.
[21] *D. Zagier*, On the values at negative integers of the zeta-function of a real quadratic field, L'Ens. math. **22** (1976), 55—95.

Department of Mathematics, Brown University, Providence, RI 02912, USA

Max-Planck-Institut für Mathematik, Gottfried-Claren-Straße 26, D-5300 Bonn 3

Eingegangen 3. September 1984