

ELLIPTIC CURVES OVER A FINITE FIELD AND THE TRACE FORMULA

NATHAN KAPLAN AND IAN PETROW

Dedicated to Professor B.J. Birch on his 85th birthday

ABSTRACT. We prove formulas for power moments for point counts of elliptic curves over a finite field k such that the groups of k -points of the curves contain a chosen subgroup. These formulas express the moments in terms of traces of Hecke operators for certain congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$. As our main technical input we prove an Eichler-Selberg trace formula for a family of congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ which include as special cases the groups $\Gamma_1(N)$ and $\Gamma(N)$. Our formulas generalize results of Birch and Ihara (the case of the trivial subgroup, and the full modular group), and previous work of the authors (the subgroups $\mathbb{Z}/2\mathbb{Z}$ and $(\mathbb{Z}/2\mathbb{Z})^2$ and congruence subgroups $\Gamma_0(2), \Gamma_0(4)$). We use these formulas to answer statistical questions about point counts for elliptic curves over a fixed finite field, generalizing results of Vlăduț, Gekeler, Howe, and others.

1. INTRODUCTION

Let E be an elliptic curve defined over a finite field \mathbb{F}_q with q elements. In this paper we always consider elliptic curves up to \mathbb{F}_q -isomorphism and whenever speaking of an elliptic curve E we always implicitly mean the isomorphism class of E . With this convention in mind, let $\mathcal{C} = \{E/\mathbb{F}_q\}$. The finite set \mathcal{C} is a probability space where a singleton $\{E\}$ occurs with probability

$$\mathbb{P}_q(\{E\}) = \frac{1}{q \# \mathrm{Aut}_{\mathbb{F}_q}(E)}.$$

Let $t_E \in \mathbb{Z}$ denote the trace of the Frobenius endomorphism associated to E . We have $t_E = q + 1 - \#E(\mathbb{F}_q)$ and by Hasse's Theorem $t_E^2 \leq 4q$. For a non-negative integer R , we consider

$$\mathbb{E}_q(t_E^{2R}) = \frac{1}{q} \sum_{E \in \mathcal{C}} \frac{t_E^{2R}}{\# \mathrm{Aut}_{\mathbb{F}_q}(E)}.$$

Birch [2, equation (4)] gave the following explicit formulas for $\mathbb{E}_q(t_E^{2R})$.

2010 *Mathematics Subject Classification.* 11G20, 11F72, 14G15 (Primary); 11F25, 14H52 (Secondary).

Theorem 1 (Birch). *For prime $p \geq 5$ we have*

$$\begin{aligned}
p\mathbb{E}_p(1) &= p \\
p\mathbb{E}_p(t_E^2) &= p^2 - 1 \\
p\mathbb{E}_p(t_E^4) &= 2p^3 - 3p - 1 \\
p\mathbb{E}_p(t_E^6) &= 5p^4 - 9p^2 - 5p - 1 \\
p\mathbb{E}_p(t_E^8) &= 14p^5 - 28p^3 - 20p^2 - 7p - 1 \\
p\mathbb{E}_p(t_E^{10}) &= 42p^6 - 90p^4 - 75p^3 - 35p^2 - 9p - 1 - \tau(p),
\end{aligned}$$

where $\tau(p)$ is Ramanujan's τ -function.

Pairing a curve with its quadratic twist shows that for any $R \geq 0$, $\mathbb{E}_q(t_E^{2R+1}) = 0$.

To state the general formula for $\mathbb{E}_q(t_E^{2R})$ we introduce some more notation. For $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ a congruence subgroup we write $S_k(\Gamma, \chi)$ for the \mathbb{C} -vector space of classical holomorphic weight k cusp forms of nebentype character χ for the group Γ , and for T a linear transformation of this vector space we write $\mathrm{Tr}(T|S_k(\Gamma, \chi))$ for its trace. If the nebentype character is trivial we may omit it from the notation. Let T_m be the m^{th} Hecke operator acting on such a space of cusp forms (see [33, Ch. 3] or [9, Ch. 5]). Let

$$a_{R,j} \stackrel{\mathrm{def}}{=} \frac{2R - 2j + 1}{2R + 1} \binom{2R + 1}{j} = \binom{2R}{j} - \binom{2R}{j - 1}.$$

For $a, b \in \mathbb{Z}$ denote the indicator function of $a = b$ by $\delta(a, b)$, and for $c \in \mathbb{N}$ denote the indicator function of the congruence $a \equiv b \pmod{c}$ by $\delta_c(a, b)$. If $q = p^v$ is a prime power (we allow $v = 0$ in which case T_1 is the identity operator) we define

$$\rho(q, k) \stackrel{\mathrm{def}}{=} -\mathrm{Tr}(T_q|S_k(\mathrm{SL}_2(\mathbb{Z}))) + \frac{k-1}{12} q^{k/2-1} \delta_2(v, 0) - \frac{1}{2} \sum_{0 \leq i \leq v} \min(p^i, p^{v-i})^{k-1} + \sigma(q) \delta(k, 2),$$

and $\rho(p^{-1}, k) = 0$. In the prime field case, the following formula is [2, equation (4)]. The general finite field case is implicit in the work of Ihara [15], see also [18].

Theorem 2 (Birch, Ihara). *For all $R \geq 0$ and $q = p^v$ with p prime, $\mathbb{E}_q(t_E^{2R+1}) = 0$ and*

$$\begin{aligned}
\mathbb{E}_q(t_E^{2R}) &= \sum_{j=0}^R a_{R,j} q^{j-1} \left(\rho(q, 2R - 2j + 2) - p^{2R-2j+1} \rho(q/p^2, 2R - 2j + 2) \right) \\
&\quad + \frac{p-1}{12q} (4q)^R \delta_2(v, 0).
\end{aligned}$$

In particular, as $q \rightarrow \infty$ we have

$$\mathbb{E}_q(t_E^{2R}) \sim C_R q^R,$$

where $C_R = \frac{1}{R+1} \binom{2R}{R}$ is the R^{th} Catalan number.

The constants C_R match the moments of the Sato-Tate distribution, and by Carleman's condition since the moments do not grow too fast they determine the limiting probability distribution of the set $\{t_E/(2\sqrt{q}) \mid E/\mathbb{F}_q\} \subset [-1, 1]$ as q tends to infinity (see e.g. [23, p. 126]).

In this paper we give a generalization of the theorems of Birch and Ihara where $\mathrm{SL}_2(\mathbb{Z})$ is replaced with a congruence subgroup, and where we count only elliptic curves over a finite field whose group of rational points contains a subgroup isomorphic to a specified group. Specifically, let A denote a finite abelian group and let Φ_A be the function defined on \mathcal{C} by

$$\Phi_A(E) = \begin{cases} 1 & \text{if there exists an injective homomorphism } A \hookrightarrow E(\mathbb{F}_q) \\ 0 & \text{otherwise.} \end{cases}$$

The main result of this paper, Theorem 3, is a generalization of Theorem 2 to the expectations

$$\mathbb{E}_q(t^R \Phi_A) = \frac{1}{q} \sum_{\substack{E \in \mathcal{C} \\ A \hookrightarrow E(\mathbb{F}_q)}} \frac{t_E^R}{\#\mathrm{Aut}_{\mathbb{F}_q}(E)}.$$

The added flexibility of the function Φ_A opens up a host of applications of our Theorem 3. For example, we give an asymptotic formula for the average exponent (also called the first invariant factor) of $E(\mathbb{F}_q)$ over \mathcal{C} . We discuss this and other applications in Section 2.

Before setting up the notation necessary to state Theorem 3 in full generality, we give two representative special cases.

Example 1. *Suppose that $q = p$ and ℓ are both prime with $\ell \neq p$. Suppose that $A = \mathbb{Z}/\ell\mathbb{Z}$. If $p \not\equiv 1 \pmod{\ell}$ then*

$$p\mathbb{E}_p(t^{2R}\Phi_A) = \frac{1}{\ell-1}C_R(p+1)p^R - \sum_{j=0}^R a_{R,j}p^j \left(\frac{\mathrm{Tr}(T_p|S_{2R-2j+2}(\Gamma_1(\ell)))}{\ell-1} + \begin{cases} 1 & p \equiv -1 \pmod{\ell} \\ 1/2 & p \not\equiv -1 \pmod{\ell} \end{cases} \right),$$

and if $p \equiv 1 \pmod{\ell}$ then

$$p\mathbb{E}_p(t^{2R}\Phi_A) = \frac{\ell}{\ell^2-1}C_R(p+1)p^R - \frac{1}{\ell-1} \sum_{j=0}^R a_{R,j}p^j \left(\mathrm{Tr}(T_p|S_{2R-2j+2}(\Gamma_1(\ell))) - \frac{1}{\ell+1} \mathrm{Tr}(T_p|S_{2R-2j+2}(\Gamma(\ell))) + \frac{1}{4}(3+(-1)^\ell) \right).$$

The leading term here when $R = 0$ gives the probability that an elliptic curve E/\mathbb{F}_p has $\#E(\mathbb{F}_p)$ divisible by a prime ℓ , a result originally due to Lenstra [25, Proposition 1.14].

Example 2. *Suppose that $q = p$ and ℓ are both prime with $\ell \neq p$. Suppose that $A = \mathbb{Z}/\ell\mathbb{Z} \times \mathbb{Z}/\ell\mathbb{Z}$. If $p \not\equiv 1 \pmod{\ell}$ then $\mathbb{E}_p(t^{2R}\Phi_A) = 0$, and if $p \equiv 1 \pmod{\ell}$ then*

$$p\mathbb{E}_p(t^{2R}\Phi_A) = \frac{1}{\ell(\ell^2-1)}C_R(p+1)p^R - \frac{1}{\ell(\ell^2-1)} \sum_{j=0}^R a_{R,j}p^j \left(\mathrm{Tr}(T_p|S_{2R-2j+2}(\Gamma(\ell))) + \frac{1}{4}(\ell^2-1)(3+(-1)^\ell) \right).$$

For E/\mathbb{F}_q and $(n, q) = 1$, $E[n] \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ over $\overline{\mathbb{F}_q}$. With respect to a $\mathbb{Z}/n\mathbb{Z}$ -module basis of $E[n]$, the action of $\mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$ gives a matrix $F \in \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ satisfying $\det(F) \equiv q$

$(\text{mod } n)$ and $\text{Tr}(F) \equiv t_E \pmod{n}$. We see that $\mathbb{Z}/n\mathbb{Z} \hookrightarrow E(\mathbb{F}_q)$ if and only if 1 is an eigenvalue of F , and that $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \hookrightarrow E(\mathbb{F}_q)$ if and only if F is the 2×2 identity matrix.

The leading constants in Example 1, $1/(\ell - 1)$ and $\ell/(\ell^2 - 1)$, have group-theoretic interpretations as the density in the coset $\{g \in \text{GL}_2(\mathbb{F}_\ell) : \det(g) \equiv p \pmod{\ell}\} \subset \text{GL}_2(\mathbb{F}_\ell)$ of the matrices with 1 as an eigenvalue, where the cases are distinguished by whether $p \equiv 1 \pmod{\ell}$ or not. Similarly, the leading constant $1/(\ell(\ell^2 - 1))$ in Example 2 is equal to the density of the identity matrix in this coset (recall that we assume $p \equiv 1 \pmod{\ell}$ in this case). We explain a link to the Chebotarev density theorem in Section 1.2.

1.1. Statement of the Main Result. The set of functions $\{t^R\}_{R \geq 0}$ is not the most natural basis of continuous functions on the interval $[-1, 1]$ in our situation. Instead, we consider the space $L^2([-1, 1], d\mu_\infty)$, where μ_∞ is the Sato-Tate measure on $[-1, 1]$ given by

$$d\mu_\infty = \frac{2}{\pi} \sqrt{1 - t^2} dt.$$

This L^2 space admits a natural orthonormal basis of polynomials, called *Chebyshev polynomials of the second kind*. For $j \geq 0$ these are defined as

$$\begin{aligned} U_0(t) &= 1 \\ U_1(t) &= 2t \\ U_{j+1}(t) &= 2tU_j(t) - U_{j-1}(t). \end{aligned}$$

The Chebyshev polynomials are particularly natural from the monodromy point of view (see Section 1.2), as we now explain. The underlying group from the monodromy perspective is SU_2 . This group has a standard 2-dimensional representation, which we denote by Std , and has a unique finite dimensional irreducible representation Sym^j of dimension $j + 1$ for each non-negative integer j given by the j^{th} symmetric power of Std . Let χ_j denote the character of the representation Sym^j . We have that $U_j(\cos \theta) = \chi_j(X_\theta)$ where X_θ is the conjugacy class in SU_2 that has eigenvalues $\{e^{i\theta}, e^{-i\theta}\}$.

In this paper we define *normalized Chebyshev polynomials* to be

$$(1) \quad U_{k-2}(t, q) = q^{k/2-1} U_{k-2} \left(\frac{t}{2\sqrt{q}} \right) = \frac{\alpha^{k-1} - \bar{\alpha}^{k-1}}{\alpha - \bar{\alpha}} \in \mathbb{Z}[q, t],$$

where $\alpha, \bar{\alpha}$ are the two roots in \mathbb{C} of $X^2 - tX + q = 0$. Let

$$(2) \quad c_{R,j} = \begin{cases} a_{R/2,j} & \text{if } R \text{ even} \\ a_{\frac{R-1}{2},j} + a_{\frac{R-1}{2},j-1} & \text{if } R \text{ odd} \end{cases}$$

be the *Chebyshev coefficients*. By an induction argument we have

$$(3) \quad t^R = \sum_{j=0}^{\lfloor R/2 \rfloor} c_{R,j} q^j U_{R-2j}(t, q).$$

The formula (3) gives a dictionary between statements about the moments $\mathbb{E}_q(t_E^R \Phi_A)$ and the averages of Chebyshev polynomials $\mathbb{E}_q(U_{k-2}(t_E, q) \Phi_A)$, so it suffices to study the latter.

The Chebyshev coefficients $c_{R,j}$ also have an interpretation in terms of the representation theory of SU_2 . We decompose $\text{Std}^{\otimes k}$ into irreducibles:

$$\text{Std}^{\otimes k} = \bigoplus_{\substack{0 \leq j \leq k \\ j \equiv k \pmod{2}}} \langle \text{Sym}^j, \text{Std}^{\otimes k} \rangle \text{Sym}^j.$$

The multiplicity coefficients $\langle \text{Sym}^j, \text{Std}^{\otimes k} \rangle$ are, up to a change of variable, the Chebyshev coefficients defined in (2). Precisely, we have $c_{R,j} = \langle \text{Sym}^{R-2j}, \text{Std}^{\otimes R} \rangle$.

In Section 4 we prove a version of the Eichler-Selberg trace formula (see Theorem 9) for the following congruence subgroups. For positive integers $M \mid N$ let

$$(4) \quad \Gamma(N, M) \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{Z}) \text{ s.t. } a, d \equiv 1 \pmod{N}, c \equiv 0 \pmod{NM} \right\}.$$

Note, for example, that $\Gamma(N, 1) = \Gamma_1(N)$ and $\Gamma(N, N) \cong \Gamma(N)$ via conjugation by $\begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}$. The Eichler-Selberg trace formula for $\text{SL}_2(\mathbb{Z})$ appears in Selberg's original paper on the trace formula [30], the generalization to $\Gamma_0(N), \chi$ under the assumption that the index of the Hecke operator is relatively prime to N was given by Hijikata [12], and the general case was achieved by Oesterlé [26] (see the paper of Cohen [4] for a description).

Let $\psi(n) = [\text{SL}_2(\mathbb{Z}) : \Gamma_0(n)] = n \prod_{p|n} (1 + 1/p)$, $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times| = n \prod_{p|n} (1 - 1/p)$, and $\phi(n) = n \prod_{p|n} (-\varphi(p))$, where the products run over primes. In particular, the function $\phi(n)$ has been chosen to be the Dirichlet convolution inverse to the function $\varphi(n^2)$. For a positive integer n and prime p we write $v_p(n)$ for the p -adic valuation of n , and we write (a, b) for the greatest common divisor of a and b . For $d \in (\mathbb{Z}/n_1\mathbb{Z})^\times$ let $\langle d \rangle$ denote the d^{th} diamond operator acting on $S_k(\Gamma(n_1, n_2))$. The operators $\langle d \rangle$ and T_q with $(q, N) = 1$ are normal and pairwise commute. For the definitions of T_q and $\langle d \rangle$ see Section 4.

We define a function $T_{n_1, \lambda}(q, d)$, which is motivated by the trace formula, as follows. We write $q = p^v$ where p is prime and v is a non-negative integer. For $\lambda \mid (d^2q - 1, n_1)$ (which is well-defined even though d is not an integer) let

$$T_{n_1, \lambda}(q, d) = \frac{\psi(n_1^2/\lambda^2)\varphi(n_1/\lambda)}{\psi(n_1^2)} (-T_{\text{trace}} + T_{\text{id}} - T_{\text{hyp}} + T_{\text{dual}}),$$

with

$$T_{\text{trace}} = \frac{1}{\varphi(n_1)} \text{Tr}(T_q \langle d \rangle | S_k(\Gamma(n_1, \lambda))),$$

$$T_{\text{id}} = \frac{k-1}{24} q^{k/2-1} \psi(n_1 \lambda) \left(\delta_{n_1}(q^{1/2}, d^{-1}) + (-1)^k \delta_{n_1}(q^{1/2}, -d^{-1}) \right),$$

$$T_{\text{hyp}} = \frac{1}{4} \sum_{i=0}^v \min(p^i, p^{v-i})^{k-1} \sum_{\substack{\tau | n_1 \lambda \\ g | p^i - p^{v-i}}} \frac{\varphi(g)\varphi(n_1(\lambda, g)/g)}{\varphi(n_1)} \\ \times \left(\delta_{n_1(\lambda, g)/g}(y_i, d^{-1}) + (-1)^k \delta_{n_1(\lambda, g)/g}(y_i, -d^{-1}) \right),$$

$$T_{\text{dual}} = \frac{\sigma(q)}{\varphi(n_1)} \delta(k, 2),$$

and where in the expressions above:

- if q is not a square then $\delta_{n_1}(q^{1/2}, \pm d^{-1}) = 0$,

- $g = (\tau, n_1\lambda/\tau)$,
- y_i is the unique element of $(\mathbb{Z}/(n_1\lambda/g)\mathbb{Z})^\times$ such that $y_i \equiv p^i \pmod{\tau}$ and $y_i \equiv p^{v-i} \pmod{n_1\lambda/\tau}$,

We also define $T_{n_1,\lambda}(p^{-1}, d) = 0$.

The main result of this paper is the following. For a finite abelian group A , let $n_1 = n_1(A)$ and $n_2 = n_2(A)$ be its first and second invariant factors, respectively. That is to say, n_1 is the largest order of a cyclic subgroup of A .

Theorem 3. *Let A be a finite abelian group of rank at most 2. Suppose that $(q, |A|) = 1$ and $k \geq 2$. If $q \equiv 1 \pmod{n_2(A)}$ we have*

$$\begin{aligned} \mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) &= \frac{1}{q\varphi(n_1/n_2)} \sum_{\nu | \frac{(q-1, n_1)}{n_2}} \phi(\nu) (T_{n_1, n_2\nu}(q, 1) - p^{k-1}T_{n_1, n_2\nu}(q/p^2, p)) \\ &\quad + q^{k/2-1} \frac{(p-1)(k-1)}{24q} (\delta_{n_1}(q^{1/2}, 1) + (-1)^k \delta_{n_1}(q^{1/2}, -1)) \end{aligned}$$

and if $q \not\equiv 1 \pmod{n_2(A)}$ then $\mathbb{E}_q(U_{k-2}(t, q)\Phi_A) = 0$.

The following special case gives the flavor of the general formula.

Example 3. *Suppose that $q = p$ and $\ell \neq p$ are primes. Suppose that A is a finite abelian group with rank at most 2. Also suppose that $p \equiv 1 \pmod{n_2(A)}$ and $(p-1, n_1(A)) = n_2(A)$. Then*

$$\begin{aligned} \mathbb{E}_p(U_{k-2}(t_E, p)\Phi_A) &= \frac{\psi(n_1^2/n_2^2)}{p\psi(n_1^2)\varphi(n_1)} \left((p+1)\delta(k, 2) - \text{Tr}(T_p | S_k(\Gamma(n_1, n_2))) \right) \\ &\quad - \frac{1}{4} \left(1 + (-1)^k \delta(n_1, 2) + \delta_{n_1}(p, 1) + (-1)^k \delta_{n_1}(p, -1) \right) \varphi(n_1) \sum_{\substack{\tau | n_1 n_2 \\ (\tau, \frac{n_1 n_2}{\tau}) | n_2}} \varphi\left(\left(\tau, \frac{n_1 n_2}{\tau}\right)\right). \end{aligned}$$

Remarks:

- (1) The formula in Theorem 3 looks complicated but it is quite usable. As $q \rightarrow \infty$ there are very few terms on the right hand side of this formula as compared to the left hand side. Of these terms, only the “trace” term is mysterious, as the “identity”, “hyperbolic”, and “dual” terms are all given in terms of straightforward arithmetic functions.

One knows quite a lot about the spaces of classical cusp forms $S_k(\Gamma(n_1, n_2))$, hence about the traces appearing in Theorem 3. This gives a lot of information that is inaccessible from the starting formula on the left hand side. As a first example, one can use Deligne’s bound on individual Hecke eigenvalues and an estimate for the dimension of $S_k(\Gamma(n_1, n_2))$ to obtain estimates as $q \rightarrow \infty$ on the left hand side that are significantly better than the trivial upper bound of $\ll q^{k/2-1}$. We explain this and several other corollaries of Theorem 3 in greater detail in Section 2.

- (2) As previously remarked, the special case $A = 1$ (the trivial group) of Theorem 3 goes back to work of Birch [2] and Ihara [15]. The cases $A = \mathbb{Z}/2\mathbb{Z}$ and $A = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ were previous work of the authors in [18], and the case of $A = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ over

the prime field \mathbb{F}_p in weight $k = 2$ was given by Kowalski in terms of \mathbb{F}_p -points on the modular curve $Y(n)$ [24, Section 6.2].

- (3) Theorem 3 requires the hypothesis that $(q, |A|) = 1$. We are cautiously optimistic that the methods of this paper could be adapted to treat the case $p \mid |A|$ as well. However, doing so would require working out in parallel to the already lengthy computations in Sections 3, 4, and 5 the special case where A is a p -group. To avoid dividing our argument into additional cases we suppose that $p \nmid |A|$ in this paper.
- (4) As was noted in [24, Section 6.2], the hyperbolic terms T_{hyp} in Theorem 3 may have an interpretation in terms of the cusps on the modular curve $X(n_1, n_2)$ corresponding to the group $\Gamma(n_1, n_2)$. For example, in Example 2 the “hyperbolic terms” are

$$\frac{1}{4} (3 + (-1)^\ell) \frac{\ell^2 - 1}{\ell(\ell^2 - 1)},$$

which is the number of cusps of $\Gamma(\ell)$ (see e.g. [9, Section 3.8]), divided by $|\text{SL}_2(\mathbb{F}_\ell)|$.

- (5) Theorem 3 could presumably be obtained working directly with modular curves, using Deligne’s equidistribution theorem and the Chebotarev density theorem (see, e.g. Katz-Sarnak [20, Chapter 9]). For example, Howe obtains the main term for a similar counting problem in this way [13]. Our results generalize Howe’s in that Theorem 3 is an explicit (not only asymptotic) formula in terms of traces of Hecke operators, and also in that we give formulas for all Chebyshev polynomials (not only U_0). We give a few more details on the geometric approach in Section 1.2. In Section 2.2 and in Theorem 4 we give applications of these formulas for $k \geq 3$, and in Section 2.4 we discuss applications of the explicit formulas for the “error terms”.
- (6) On the other hand, our approach has some considerable advantages over the monodromy approach:
- The natural geometric setting is probably about algebraic stacks (note the automorphisms in the probability function \mathbb{P}_q) and the analytic approach here hides the complications involved in that theory. Dealing with the primes 2 and 3 geometrically might be quite involved, but our approach is not much more difficult for these primes.
 - Any geometric proof that could reproduce the exact formulas in Theorem 3 would necessarily be quite delicate and complicated.
 - It is by no means clear that a geometric proof dealing only with the asymptotic order of magnitude would be as strong quantitatively in its applications, i.e. that the error term in Corollary 1 below would be as uniform in n_1 and n_2 .

1.2. Alternate Approach. In this section we sketch an alternate approach to some of our results using geometric techniques and monodromy computations. This approach naturally explains some features of the computations that follow in the rest of the paper, and extends (at least to asymptotic order) to variants of the problem where it is not clear a trace formula approach would work.

Consider any one-parameter family of elliptic curves (E_λ) defined over \mathbb{F}_q and consider those λ such that $E_\lambda(\mathbb{F}_q)$ contains a subgroup isomorphic to A . The following proposition describes the special case where (E_λ) is the Legendre family. It was suggested by an anonymous referee.

Proposition 1. *Let $\mathcal{E} \rightarrow \mathbb{A}^1 - \{0, 1\}$ be the family of elliptic curves over $\mathbb{Z}[1/2]$ given by the Legendre family*

$$E_\lambda : y^2 = x(x-1)(x-\lambda).$$

For a finite field \mathbb{F}_q with $(q, 2) = 1$ and any $\lambda \in \mathbb{F}_q - \{0, 1\}$, let $t(\lambda)$ be the corresponding trace of Frobenius. Let A be a fixed finite abelian group of rank at most 2 and with odd order. As $q \rightarrow \infty$ through any sequence with $(q, 2|A|) = 1$ and $q \equiv 1 \pmod{n_2(A)}$ the finite sets

$$\left\{ \frac{t(\lambda)}{\sqrt{q}} : \lambda \in \mathbb{F}_q - \{0, 1\} \text{ and } A \hookrightarrow E_\lambda(\mathbb{F}_q) \right\}$$

become equidistributed with respect to the Sato-Tate measure.

Proof Sketch. Let E be an elliptic curve defined over \mathbb{F}_q . Let A be a fixed finite abelian group of rank at most 2, and let $n_1 = n_1(A)$ and $n_2(A) = n_2$ be its first and second invariant factors. Let $n = n_1 n_2$. The condition that $A \hookrightarrow E(\mathbb{F}_q)$ can be detected via the $\text{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$ -action on $E[n]$. Specifically, this action determines a well-defined conjugacy class $(\text{frob}_q) \subset \text{GL}_2(\mathbb{Z}/n\mathbb{Z})$, and $A \hookrightarrow E(\mathbb{F}_q)$ if and only if (frob_q) has one eigenvalue that is 1 modulo n_1 and (frob_q) is trivial modulo n_2 . We say a conjugacy class of $\text{GL}_2(\mathbb{Z}/n\mathbb{Z})$ is of *type A* if these conditions are satisfied.

Let $K = \mathbb{F}_q(t)$, $G_{\text{ar}} = \text{Gal}(K^{\text{sep}}/K)$ and $G_{\text{geo}} = \text{Gal}(K^{\text{sep}}/K.\overline{\mathbb{F}_q})$ be the arithmetic and geometric Galois groups of K . To show that the distribution of $t(\lambda)$ is independent of the Galois action on n -torsion points, we fix a prime $\ell \nmid q$ and consider two Galois representations:

(1)

$$\rho : G_{\text{ar}} \rightarrow \text{GL}(V),$$

given by the ℓ -adic sheaf $\mathcal{F} = R^1 \pi_* \mathbb{Q}_\ell(1/2)$ lisse outside of $\{0, 1, \infty\}$ (see e.g. [20, Section 9.1.11]) where $\pi : \mathcal{E} \rightarrow \mathbb{P}^1$, and V is the 2-dimensional \mathbb{Q}_ℓ -vector space given intrinsically by the stalk $\mathcal{F}_{\overline{\eta}}$ for $\overline{\eta}$ a geometric generic point of \mathbb{P}^1 .

(2)

$$\rho_n : G_{\text{ar}} \rightarrow \text{GL}_2(\mathbb{Z}/n\mathbb{Z})$$

given by the Galois action on n -torsion.

Let $U \subset \mathbb{P}^1$ be the affine open of \mathbb{P}^1 avoiding $\{0, 1, \infty\}$. For each $\lambda \in U(\mathbb{F}_q)$ we have inertia and decomposition groups $I_\lambda \trianglelefteq D_\lambda \leq G_{\text{ar}}$ and an element $\text{frob}_{\lambda, q} \in D_\lambda/I_\lambda \simeq \text{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$, which is the inverse image of the geometric Frobenius. Therefore for each $\lambda \in U(\mathbb{F}_q)$ we have a well-defined conjugacy class $\text{frob}_{\lambda, q}$ in a quotient of G_{ar} (the conjugacy class $\text{frob}_{\lambda, q}$ in G_{ar} itself depends on a choice of lift modulo I_λ). We have in particular for all $\lambda \in U(\mathbb{F}_q)$ that $\text{Tr} \rho(\text{frob}_{\lambda, q}) = -t(\lambda)/\sqrt{q}$, and that $\rho_n(\text{frob}_{\lambda, q}) = (\text{frob}_q)$, the conjugacy class of $\text{GL}_2(\mathbb{Z}/n\mathbb{Z})$ defined via the curve E_λ in the first paragraph of this proof.

By Lemma 1 (see Section 2.2), our goal is to show that for each $j \geq 1$ we have

$$\frac{\sum_{\lambda \in U(\mathbb{F}_q)} U_j(t(\lambda)/2\sqrt{q}) \Phi_A(E_\lambda)}{\sum_{\lambda \in U(\mathbb{F}_q)} \Phi_A(E_\lambda)} \sim 0$$

as $q \rightarrow \infty$. To do this, we use the Lefschetz trace formula and the Riemann hypothesis of Deligne to estimate the quantity in the numerator.

We begin by calculating the monodromy groups of ρ and ρ_n . By theorems of Deligne and because of the normalization by the Tate twist we have (see [20, Section 10.1.16]) that

$$\overline{\iota \rho(G_{\text{geo}})} = \overline{\iota \rho(G_{\text{ar}})} = \text{SL}_2(\mathbb{C})$$

where $\iota : \mathbb{Q}_\ell \rightarrow \mathbb{C}$ is a fixed complex embedding and the bar denotes the Zariski closure. We also let $G_n = \rho_n(G_{\text{ar}})$ and $G_{n,0} = \rho_n(G_{\text{geo}}) \trianglelefteq G_n$. The geometric monodromy group H of $\rho \times \rho_n$ is contained in $\text{SL}_2 \times G_{n,0}$. We claim it is equal to this group. The connected component of the identity H^0 is of the same dimension as $\text{SL}_2(\mathbb{C})$ so equals $\text{SL}_2(\mathbb{C}) \times \{1\}$. Since H also surjects onto $G_{n,0}$ we conclude that $H = \text{SL}_2(\mathbb{C}) \times G_{n,0}$.

Let Λ be any irreducible representation of $\text{SL}_2(\mathbb{C})$ and π any irreducible representation of G_n . If $\Lambda \otimes \pi$ is not trivial on the geometric monodromy group of $\rho \times \rho_n$ we apply the Grothendieck-Lefschetz trace formula and the Riemann hypothesis of Deligne (see e.g. [20, 9.2.6(2) and 9.2.6(3)]) to deduce

$$(5) \quad \sum_{\lambda \in U(\mathbb{F}_q)} \text{Tr}(\Lambda(\rho(\text{frob}_{\lambda,q}))) \text{Tr}(\pi(\rho_n(\text{frob}_{\lambda,q}))) \ll_{\Lambda,n} \sqrt{q}.$$

Here the implied constant depends on n and Λ but it can be shown using the Grothendieck-Ogg-Safarevich formula that this constant is $\ll |G_n| \dim(\Lambda) \dim(\pi) \ll |G_n|^{3/2} \dim(\Lambda)$, with an absolute implied constant.

Choosing Λ to be the j^{th} symmetric power representation, we get $\text{Tr}(\Lambda(\rho(\text{frob}_{\lambda,q}))) = U_j(t(\lambda)/2\sqrt{q})$. By the first paragraph of the proof, we expand $\Phi_A(E_\lambda)$ spectrally into finite dimensional irreducible representations of G_n :

$$\Phi_A(E_\lambda) = \begin{cases} 1 & \text{if } \rho_n(\text{frob}_{\lambda,q}) \text{ is of type } A \\ 0 & \text{else} \end{cases} = \sum_{\pi} c(A, \pi) \text{Tr}(\pi(\rho_n(\text{frob}_{\lambda,q}))).$$

We then have by (5) that

$$\sum_{\lambda \in U(\mathbb{F}_q)} U_j(t(\lambda)/2\sqrt{q}) \Phi_A(E_\lambda) \ll_{A,j} q^{-1/2}.$$

The Chebotarev density theorem over finite fields (see [20, Sections 9.7.10 and 9.7.11]) implies that

$$(6) \quad \sum_{\lambda \in U(\mathbb{F}_q)} \Phi_A(E_\lambda) = (q-2) \frac{|\{g \in G_n : \bar{g} \text{ has type } A\}|}{|G_{n,0}|} + O_n(\sqrt{q}),$$

where \bar{g} is the canonical projection $G_n \rightarrow G_n/G_{n,0}$. To show equidistribution with respect to the Sato-Tate measure, it therefore suffices to show that the numerator in (6) is non-zero.

A result of Igusa [14, Theorem 3] states that for n odd we have $G_{n,0} = \text{SL}_2(\mathbb{Z}/n\mathbb{Z})$. Therefore the cosets $G_n/G_{n,0}$ correspond to determinants of elements, and as $\det(\text{frob}_{\lambda,q}) = q^{-1}$ we have that the matrix

$$\begin{pmatrix} q & 0 \\ 0 & 1 \end{pmatrix} \in G_n \subseteq \text{GL}_2(\mathbb{Z}/n\mathbb{Z})$$

has one eigenvalue equal to 1 and is the identity modulo n_2 , by our assumption that $q \equiv 1 \pmod{n_2(A)}$. Therefore, the main term of (6) does not vanish, which completes the proof. \square

1.3. Sketch of the Proof of Theorem 3. For A an abelian group and $t \in \mathbb{Z}$ let $\mathcal{C}(A) = \{E/\mathbb{F}_q : A \hookrightarrow E(\mathbb{F}_q)\}$ and $\mathcal{C}(A, t) = \{E/\mathbb{F}_q : A \hookrightarrow E(\mathbb{F}_q) \text{ and } t_E = t\}$. These sets are empty unless A is finite of rank at most 2 and $t^2 \leq 4q$.

Step 1. The first step in the proof of Theorem 3 is to fiber $\mathcal{C}(A)$ over isogeny classes, which are parameterized by $t \in \mathbb{Z}$ with $t^2 \leq 4q$. We have

$$\begin{aligned}
 \mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) &= \frac{1}{q} \sum_{\substack{E/\mathbb{F}_q \\ A \hookrightarrow E(\mathbb{F}_q)}} \frac{U_{k-2}(t_E, q)}{\#\text{Aut}_{\mathbb{F}_q}(E)} = \sum_{t^2 \leq 4q} U_{k-2}(t, q) \left(\frac{1}{q} \sum_{\substack{t_E=t \\ A \hookrightarrow E(\mathbb{F}_q)}} \frac{1}{\#\text{Aut}_{\mathbb{F}_q}(E)} \right) \\
 (7) \qquad \qquad \qquad &= \sum_{t^2 \leq 4q} U_{k-2}(t, q) \mathbb{P}_q(\mathcal{C}(A, t)).
 \end{aligned}$$

Step 2. In Section 3, the sizes of the fibers $\mathbb{P}_q(\mathcal{C}(A, t))$ are expressed in terms of sums of ideal class numbers of orders in imaginary quadratic fields. When $A = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ such results go back to Deuring [8], Lenstra [25], Schoof [29], and Waterhouse [35], and are stated as Lemmas 2, 3 and 4 below. We extend these results to general rank at most 2 finite abelian groups in Proposition 2. To state these results precisely, we introduce in (32) certain class numbers $H_{n_1, n_2}(t, q, d)$ where $n_1 = n_1(A)$ and $n_2 = n_2(A)$.

Step 3. In Section 4 we prove a version of the Eichler-Selberg trace formula for the groups $\Gamma(n_1, n_2)$ by summing the formula for $\Gamma_0(n), \chi$ over nebentype characters χ . The computation is very explicit and delicate, but in the end it turns out that the elliptic term involves a similar sum over the same class numbers $H_{n_1, n_2}(t, q, d)$. The full formula is given in Theorem 9.

Step 4. The sum over isogeny classes t in (7) is not exactly the same as the elliptic terms in the trace formula (Theorem 9), but they are close enough that we can compare them explicitly in Section 5. After some manipulation, the trace formula captures all of the elliptic curves in (7) where the ring of endomorphisms *over the base field* is an order in a imaginary quadratic field, and not only the ordinary curves. The isogeny classes corresponding to $t^2 = 4q$ are the only ones where the endomorphism rings over the base field are orders in quaternion algebras, and these have to be considered separately in the proof of Theorem 3. We remark also that the comparison in Section 5 is particularly simple in the case that \mathbb{F}_q is a prime field with $p \geq 5$.

1.4. Index of Notation. In this section we give an index of notation used throughout the paper. If definitions are easy to state, we give them here, but for more involved notation we give only a location for where a full definition can be found.

Notation	Definition	Location
q	A prime power	Page 1
\mathbb{F}_q	A finite field with q elements	Page 1
p	The characteristic of \mathbb{F}_q	Page 2
ℓ	A prime $\ell \neq p$	Page 3
t	An integer satisfying $t^2 \leq 4q$	Page 9
A	Finite abelian group of rank at most 2	Page 3
\prod_p, \prod_ℓ	Products over primes	Page 5
$\sigma(n)$	Sum of divisors function: $\sigma(n) = \sum_{d n} d$	Page 2
$\varphi(n)$	Euler phi function: $\varphi(n) = n \prod_{p n} (1 - 1/p)$	Page 5
$\psi(n)$	$n \prod_{p n} (1 + 1/p)$	Page 5
$\phi(n)$	$n \prod_{p n} (-\varphi(p))$	Page 5
$\mu(n)$	Möbius function	Page 19
$v_p(n)$	p -adic valuation of n	Page 5
(a, b)	Greatest common divisor of a and b	Page 5
$\tau(n)$	Number of divisors of n	Page 15
$\omega(n)$	Number of distinct prime factors of n	Page 24
$\lambda(n)$	$(-1)^{\omega(n)}$	Page 24
$\lfloor x \rfloor$	The greatest integer $\leq x$	Page 14
$\delta(a, b)$	Indicator function of $a = b$	Page 2
$\delta_c(a, b)$	Indicator function of $a \equiv b \pmod{c}$	Page 2
$D(t; n)$	$\delta_n(dq + d^{-1}, t)$	Page 23
O_x, \ll_x	Big O notation. The constant may depend on x	Page 13
$a_{R,j}$	$\binom{2R}{j} - \binom{2R}{j-1}$	Page 2
$c_{R,j}$	Chebyshev coefficients	Page 4
C_R	Catalan number: $a_{R,R}$	Page 2
$U_j(t)$	Chebyshev polynomials of the second kind	Page 4
$U_{k-2}(t, q)$	Normalized Chebyshev polynomials	Page 4
$n_1(A), n_2(A)$	First and second invariant factors of A	Page 6
$n_1(E), n_2(E)$	First and second invariant factors of $E(\mathbb{F}_q)$	Page 14
E/\mathbb{F}_q	Isomorphism class of the elliptic curve E over \mathbb{F}_q	Page 1
t_E	Trace of the Frobenius endomorphism associated to E	Page 1
\mathcal{C}	Set of all isomorphism classes of elliptic curves over \mathbb{F}_q	Page 1
$\mathcal{C}(A)$	Set of all E/\mathbb{F}_q where $\Phi_A(E) = 1$	Page 9
$\mathcal{C}(A, t)$	$\{E/\mathbb{F}_q : \Phi_A(E) = 1 \text{ and } t_E = t\}$	Page 9
Φ_A	Indicator function on \mathcal{C} of $A \hookrightarrow E(\mathbb{F}_q)$	Page 2
$\mathbb{P}_q(*)$	Probability measure on \mathcal{C}	Page 1
$\mathbb{E}_q(*)$	Expectation of a random variable on \mathcal{C}	Page 1
$v(n_1, n_2)$	Constant for main term from Corollary 1	Page 13
$h(d)$	Class number of the quadratic order of discriminant d	Page 21
$h_w(d)$	Class number weighted by size of unit group	Page 21
$H(\Delta)$	Hurwitz-Kronecker class number	Page 21
$H_{n_1, n_2}(t, q, d)$	Modified class numbers	Page 24
$H_{n_1, n_2}^*(t, q, d)$	Class numbers for supersingular contribution	Page 26
$\left(\frac{a}{n}\right)$	Kronecker symbol	Page 21

Notation	Definition	Location
$\Gamma(N, M)$	Congruence subgroup of (4)	Page 5
T_m	m^{th} Hecke operating acting on a space of cusp forms	Page 28
$\langle d \rangle$	d^{th} diamond operator acting on $S_k(\Gamma(n_1, n_2))$	Page 28
$S_k(\Gamma, \chi)$	Space of weight k cusp forms of character χ for Γ	Page 2
$\text{Tr}(T S_k(\Gamma, \chi))$	Trace of T acting on the vector space $S_k(\Gamma, \chi)$	Page 29
$T_{n_1, \lambda}(q, d)$	Function motivated by the trace formula	Page 5
$\rho(q, k)$	A special case of $T_{n_1, \lambda}(q, d)$	Page 2
$\omega_A(q, d), \omega_A^*(q, d)$	Functions defined in terms of class numbers	Page 27
$\Sigma_{n_1, n_2}(q, d)$	A sum of class numbers	Page 53
T_{trace}	Part of $T_{n_1, \lambda}(q, d)$	Page 5
T_{id}	Part of $T_{n_1, \lambda}(q, d)$	Page 5
T_{hyp}	Part of $T_{n_1, \lambda}(q, d)$	Page 5
T_{dual}	Part of $T_{n_1, \lambda}(q, d)$	Page 5
$T_{\chi}^{(*)}$	Parts of Theorem 10, where $*$ = $i, e, h,$ or d	Page 31
νn_1	Full divisor: for all primes $\ell \nu$, $v_{\ell}(\nu) = v_{\ell}(n_1)$	Page 23
$a b^{\infty}$	b is divisible by each prime dividing a	Page 23
$\mu \prec \nu$	A partial order on integers, see Section 3	Page 23
$D_{\nu, \mu}(t)$	A function related to $D(t; n)$, see Section 3	Page 23
$\mu_{\chi}(t, m, q)$	See Theorem 10	Page 30
\sum^{*m}	See Theorem 10	Page 30
Δ	$t^2 - 4q$	Page 32
Δ^*	$\Delta/4$	Page 36
m	An integer such that $m^2 \Delta$	Page 32
α	An integer such that $0 \leq \alpha \leq \beta \leq \gamma$	Page 32
γ	An integer such that $0 \leq \alpha \leq \beta \leq \gamma$	Page 32
β	The integer such that $(d^2q - 1, \ell^{\gamma}) = \ell^{\beta}$	Page 32
κ	$v_{\ell}(m)$	Page 32
ν	$v_{\ell}(\Delta)$	Page 32
$W(d)$	A sum involving \sum^{*m} , see Section 5.2	Page 32
$C(t, q, d)$	Used in the proof of Theorem 9, see Section 5.2	Page 32
$c_{\kappa}(t, q, d)$	Used in the proof of Theorem 9, see Section 5.5	Page 34
$V(\tau, d)$	Used in the proof of Theorem 9, see Section 5.3	Page 33
$S(a, n)$	Number of solutions of $x^2 - a \equiv 0 \pmod{n}$	Page 36
$C_{K, N, M}(t, q, d)$	Used in the proof of Theorem 9, see Section 5.6	Page 51

2. APPLICATIONS

2.1. Points on Elliptic Curves over Finite Field Extensions. Our first application emphasizes expectations of functions of t_E . We give a generalization of the work of Brock and Granville [3] on “quadratic excess”, i.e. that there are extra points in quadratic extensions on curves over finite fields. For a geometric explanation of this phenomenon, see the paper of Katz [19].

Let E be an elliptic curve defined over a finite field \mathbb{F}_q , and \mathbb{F}_{q^r} be the degree r extension field. Then the number of \mathbb{F}_{q^r} -points on E is given by $E(\mathbb{F}_{q^r}) = q^r + 1 - (\alpha^r + \bar{\alpha}^r)$, where

$\alpha, \bar{\alpha}$ are the two roots of $X^2 - t_E X + q$. In particular, $\alpha\bar{\alpha} = q$. We have by (3) that

$$(\alpha^r + \bar{\alpha}^r) = (\alpha^r + \bar{\alpha}^r) \frac{(\alpha - \bar{\alpha})}{(\alpha - \bar{\alpha})} = \begin{cases} U_r(t, q) - qU_{r-2}(t, q) & \text{if } r \geq 2 \\ U_1(t, q) & \text{if } r = 1. \end{cases}$$

We use this to compute the average number of points on elliptic curves over extension fields. If $r \geq 2$ we have

$$\mathbb{E}_q(\#E(\mathbb{F}_{q^r})\Phi_A) = (q^r + 1)\mathbb{E}_q(\Phi_A) - \mathbb{E}_q(U_r(t_E, q)\Phi_A) + q\mathbb{E}_q(U_{r-2}(t_E, q)\Phi_A),$$

and we will see in below Section 2.2 that

$$\mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) = \begin{cases} \mathbb{E}_q(\Phi_A) + O_{A,\varepsilon}(q^{-1/2+\varepsilon}) & \text{if } k = 2 \\ O_{A,\varepsilon}\left(q^{\frac{k-3}{2}+\varepsilon}\right) & \text{if } k \geq 3. \end{cases}$$

Therefore we have

$$\mathbb{E}_q(\#E(\mathbb{F}_{q^r})\Phi_A) = \begin{cases} (q^2 + q)\mathbb{E}_q(\Phi_A) + O_{A,\varepsilon}(q^{1/2+\varepsilon}) & \text{if } r = 2 \\ q^r\mathbb{E}_q(\Phi_A) + O_{A,\varepsilon}\left(q^{\frac{r-1}{2}+\varepsilon}\right) & \text{if } r \geq 3, \end{cases}$$

recovering the Brock-Granville quadratic excess for the family of elliptic curves with $A \hookrightarrow E(\mathbb{F}_q)$.

2.2. Families of Curves over \mathbb{F}_q and the Sato-Tate Distribution. Our second application is the Sato-Tate equidistribution of traces of the Frobenius endomorphism for several families of elliptic curves over \mathbb{F}_q . Let $\mathcal{F} \subseteq \mathcal{C}$ be a subset of elliptic curves over \mathbb{F}_q and let $\Phi(\mathcal{F})$ be the indicator function of \mathcal{F} , e.g.,

$$\Phi(\mathcal{C}(A)) = \Phi_A.$$

Similarly, we often drop the \mathcal{C} from the notation in the families that we consider below. We study the equidistribution of t_E for $E \in \mathcal{F}$ via the following lemma, which is immediate from the definition found in [32, §1].

Lemma 1. *The traces of the Frobenius t_E for $E \in \mathcal{F}$ are equidistributed with respect to the Sato-Tate measure if for all $j \geq 1$ we have*

$$\lim_{q \rightarrow \infty} \frac{\mathbb{E}_q(U_j(t_E, q)\Phi(\mathcal{F}))}{q^{j/2}\mathbb{E}_q(\Phi(\mathcal{F}))} = 0.$$

For n_1, n_2 two natural numbers with $n_2 \mid n_1$ and $q \equiv 1 \pmod{n_2}$ let

$$v(n_1, n_2) = \frac{n_1}{\psi(n_1)\varphi(n_1)n_2^2} \prod_{\ell \mid \frac{n_1}{n_2}} \left(1 + \ell^{-1-2v_\ell\left(\frac{(q-1)n_1}{n_2}\right)}\right).$$

We have the trivial estimate $\mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) \ll q^{k/2-1}$. The following is the main corollary of Theorem 3.

Corollary 1. *Let A be a finite abelian group of rank at most 2 with $(q, |A|) = 1$ and $k \geq 2$ an integer. When $q \equiv 1 \pmod{n_2(A)}$ we have*

$$\mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) = v(n_1(A), n_2(A)) \left(\delta(k, 2) + O_\varepsilon(kn_2(A)n_1(A)^{2+\varepsilon}q^{\frac{k-3}{2}+\varepsilon}) \right).$$

In particular, the traces of the Frobenius t_E for $E \in \mathcal{C}(A)$ become equidistributed with respect to the Sato-Tate measure as $q \rightarrow \infty$ through prime powers $q \equiv 1 \pmod{n_2(A)}$. The equidistribution is uniform in A as soon as $q \gg n_2(A)^2 n_1(A)^{4+\delta}$ for some $\delta > 0$.

Corollary 1 follows from Theorem 3 and Deligne's bound on the Hecke eigenvalues of modular forms. For a similar calculation, see the proof of Theorem 6. One could obtain an asymptotic estimate from Theorem 3 without Deligne's work by using a bound of Selberg [31] on Hecke eigenvalues as Birch does in [2], but one would get a weaker error term.

By inclusion-exclusion arguments we can also show the distribution of t_E tends to the Sato-Tate distribution for several other families of elliptic curves. We give some examples here.

- (1) Howe computes the probability that $N \mid \#E(\mathbb{F}_q)$ as $q \rightarrow \infty$ [13, Theorem 1.1]. Let $\mathcal{C}(N \mid \#E(\mathbb{F}_q)) \subset \mathcal{C}$ denote the set of curves for which $N \mid \#E(\mathbb{F}_q)$, and let $\Phi(N \mid \#E(\mathbb{F}_q))$ be its indicator function. For a prime ℓ , let $A_\ell(a, b)$ denote the group $\mathbb{Z}/\ell^a\mathbb{Z} \times \mathbb{Z}/\ell^b\mathbb{Z}$. We see that

$$(8) \quad \Phi(N \mid \#E(\mathbb{F}_q)) = \prod_{\ell \mid N} \left(\Phi_{A_\ell(v_\ell(N), 0)} + \sum_{k=1}^{\lfloor \frac{v_\ell(N)}{2} \rfloor} \left(\Phi_{A_\ell(v_\ell(N)-k+1, k)} - \Phi_{A_\ell(v_\ell(N)-k, k)} \right) \right).$$

Expanding $U_{k-2}(t_E, q)\Phi(N \mid \#E(\mathbb{F}_q))$ using (8) and applying Corollary 1 to each term shows that for $q \rightarrow \infty$ sufficiently fast with respect to N , the t_E for $E \in \mathcal{C}(N \mid \#E(\mathbb{F}_q))$ become equidistributed with respect to the Sato-Tate measure.

- (2) Let $\mathcal{C}(\ell - \text{part}(\alpha, \beta)) \subset \mathcal{C}$ be the set of curves such that the ℓ -primary part of $E(\mathbb{F}_q)$ is isomorphic to $\mathbb{Z}/\ell^\alpha\mathbb{Z} \times \mathbb{Z}/\ell^\beta\mathbb{Z}$. Gekeler computes $\mathbb{P}_q(\mathcal{C}(\ell - \text{part}(\alpha, \beta)))$ for $q \rightarrow \infty$ through primes [10, Formula (3.9)]. Let $\Phi(\ell - \text{part}(\alpha, \beta))$ be the indicator function of $\mathcal{C}(\ell - \text{part}(\alpha, \beta))$. Then

$$(9) \quad \Phi(\ell - \text{part}(\alpha, \beta)) = \Phi(A_\ell(\alpha, \beta)) - \Phi(A_\ell(\alpha + 1, \beta)) - \Phi(A_\ell(\alpha, \beta + 1)) \\ + \Phi(A_\ell(\alpha + 1, \beta + 1))$$

when $\beta < \alpha$, and

$$(10) \quad \Phi(\ell - \text{part}(\alpha, \alpha)) = \Phi(A_\ell(\alpha, \alpha)) - \Phi(A_\ell(\alpha + 1, \alpha))$$

when $\alpha = \beta$. Applying Corollary 1 to (9) and (10), we recover formulas of Gekeler for all finite fields \mathbb{F}_q [10]. Expanding $U_{k-2}(t_E, q)\Phi(\ell - \text{part}(\alpha, \beta))$ and applying Corollary 1 to each term shows that the distribution of t_E over $E \in \mathcal{C}(\ell - \text{part}(\alpha, \beta))$ becomes equidistributed with respect to the Sato-Tate measure as $q \rightarrow \infty$ through prime powers $q \equiv 1 \pmod{\ell^\beta}$. Applying Theorem 3 to (9) and (10) gives explicit formulas in terms of traces of Hecke operators for these counts.

- (3) Let $\Phi(n_2 = m)$ be the indicator function of the family of elliptic curves $\mathcal{C}(n_2 = m) = \{E/\mathbb{F}_q : n_2(E(\mathbb{F}_q)) = m\}$. In particular, $\mathcal{C}(n_2 = 1)$ is the set of isomorphism classes of curves with cyclic group structure over \mathbb{F}_q . In Theorem 6 we give asymptotic formulas for $\mathbb{E}_q(U_{k-2}(t_E, q)\Phi(n_2 = m))$, which show by Lemma 1 that the t_E for curves $E \in \mathcal{C}(n_2 = m)$ become equidistributed with respect to the Sato-Tate measure as soon as $q \gg m^{6+\delta}$ for any $\delta > 0$.

2.3. Averages for Invariant Factors of $E(\mathbb{F}_q)$.

Theorem 4. Let $c(q)$ be defined by

$$c(q) \stackrel{\text{def}}{=} \prod_{\ell^\alpha \parallel q-1} \left(1 - \frac{1}{\ell^2(\ell+1)} \frac{1-\ell^{-4\alpha}}{1-\ell^{-4}} \right).$$

We have that

$$\mathbb{E}_q(n_1(E)) = c(q)q + O_\varepsilon(q^{1/2+\varepsilon}).$$

Note that for all q

$$.8758 < \prod_{\ell} \left(1 - \frac{\ell^2}{(\ell^4-1)(\ell+1)} \right) < c(q) \leq 1.$$

Asymptotic formulas for the higher moments $\mathbb{E}_q(n_1(E)^R)$ are also accessible by our methods.

Theorem 5. Let $b(q)$ be defined by

$$b(q) \stackrel{\text{def}}{=} \prod_{\ell^\alpha \parallel q-1} \left(1 + \frac{1}{\ell(\ell+1)} \frac{1-\ell^{-2\alpha}}{1-\ell^{-2}} \right).$$

We have that

$$\mathbb{E}_q(n_2(E)) = b(q) + \frac{pq^{-1/2}}{12} \mathbb{1}_{q=\square} + O_\varepsilon(q^{-1/3+\varepsilon}),$$

where $\mathbb{1}_{q=\square} = 1$ if q is a square and is 0 otherwise. Note that for all q

$$1 \leq b(q) < \prod_{\ell} \left(1 + \frac{\ell}{(\ell^2-1)(\ell+1)} \right) < 1.45004.$$

Theorems 4 and 5 are both consequences of the following result.

Theorem 6. If $m \nmid q-1$ then $\mathbb{E}_q(U_{k-2}(t_E, q)\Phi(n_2 = m)) = 0$. If $m \mid q-1$ then

$$\begin{aligned} \mathbb{E}_q(U_{k-2}(t_E, q)\Phi(n_2 = m)) &= \frac{(q+1)\delta(k, 2)}{q\psi(m^2)\varphi(m)} \prod_{\substack{\ell \mid \frac{q-1}{m} \\ \ell \nmid m}} \left(1 - \frac{1}{\ell(\ell^2-1)} \right) \prod_{\substack{\ell \mid \frac{q-1}{m} \\ \ell \mid m}} \left(1 - \frac{1}{\ell^3} \right) \\ &\quad + O(kq^{\frac{k-3}{2}}\tau(q-1)\log q), \end{aligned}$$

where $\tau(n)$ is the number of divisors of n , and 5 is an admissible constant in the O notation. In particular, we have that the t_E for $E \in \mathcal{C}(n_2 = m)$ become equidistributed with respect to the Sato-Tate measure as $q \rightarrow \infty$ through prime powers $q \equiv 1 \pmod{m}$. The equidistribution is uniform in m when $q \gg m^{6+\delta}$ for some $\delta > 0$.

We prove Theorems 4 and 5 in Section 2.5 and 6 in Section 2.6. Theorem 6 with $k = 2$ and $m = 1$ gives the number of \mathbb{F}_q -isomorphism classes of elliptic curves over \mathbb{F}_q with cyclic group structure, recovering a result of Vlăduț [34]. Cojocaru has shared with us a preprint [5] in which an asymptotic formula for $\mathbb{E}_q(\Phi(n_2 = m))$ is derived from the results of Howe [13].

The method of proof of Theorem 5 also allows one to compute averages of several types of arithmetic functions of $n_2(E)$ over \mathcal{C} , such as the divisor function, $\mathbb{E}_q(\tau(n_2(E)))$, or a Dirichlet character, $\mathbb{E}_q(\chi(n_2(E)))$. Such applications were suggested to us by Cojocaru at

the Arizona Winter School 2016, see the related work of her project group [1]. On the other hand, computing averages of arithmetic functions of $n_1(E)$ over \mathcal{C} seems more challenging.

2.4. Other applications. We briefly sketch some applications of our work that we do not pursue further in this paper.

Gekeler [10] studies elliptic curves not over one prime field \mathbb{F}_p , but rather takes averages over all fields \mathbb{F}_p with $p \leq X$, where $X > 1$. In this situation he applies results of Howe for individual primes [13], to prove that

$$(11) \quad \frac{|\{E/\mathbb{F}_p \mid p \leq X, E(\mathbb{F}_p)[\ell^\infty] \cong \mathbb{Z}/\ell^\alpha\mathbb{Z} \times \mathbb{Z}/\ell^\beta\mathbb{Z}\}|}{|\{E/\mathbb{F}_p \mid p \leq X\}|} = g^{(\ell)}(\alpha, \beta) + O_{\alpha, \beta, \ell}(X^{-1/2}),$$

for an explicit constant $g^{(\ell)}(\alpha, \beta)$ given in [10, Equation (2.3)]. Applying our Theorem 3 in place of the estimate for $w(m, n)$ found on page 245 of [13] gives an explicit expression for the error in term in (11) in terms of eigenvalues of Hecke operators.

Following the same steps as Section 3 of [10] we estimate (11) but now in Gekeler's step (3.11) we may exploit cancellation among the eigenvalues of T_p as p varies over $p \leq X$ to give a better error term (using e.g. Theorem 5.40 (or assuming GRH, Theorem 5.15) of [16]). We may thus improve the error in Gekeler's result to $O_{\ell, \alpha, \beta}(X^{-1/2} \exp(-C\sqrt{\log X}))$, or under GRH to $O_{\ell, \alpha, \beta}(X^{-1}(\log X)^2)$. Interestingly, a similar calculation shows that

$$(12) \quad \frac{|\{E/\mathbb{F}_{p^2} \mid p^2 \leq X, E(\mathbb{F}_{p^2})[\ell^\infty] \cong \mathbb{Z}/\ell^\alpha\mathbb{Z} \times \mathbb{Z}/\ell^\beta\mathbb{Z}\}|}{|\{E/\mathbb{F}_{p^2} \mid p^2 \leq X\}|}$$

has a lower-order main term of size asymptotic to $cX^{-1/2}$ for some $c = c_{\ell, \alpha, \beta}$ depending on ℓ, α, β which comes from the hyperbolic and supersingular terms of Theorem 3.

Next we mention briefly two standard applications of the Eichler-Selberg trace formula for $S_k(\Gamma_0(N), \chi)$ that should generalize to $S_k(\Gamma(N, M))$ using Theorem 9.

One can give simple formulas for the dimension of the space of cusp forms $\dim S_k(\Gamma_0(N), \chi)$ by studying the trace formula when $q = 1$: the Hecke operator T_1 is just the identity on a space of cusp forms, and so $\text{Tr}(T_1|S_k(\Gamma_0(N), \chi)) = \dim S_k(\Gamma_0(N), \chi)$. See for example [28, Cor 8], where Ross carefully derives this formula. Theorem 9 can be used to give a similar formula for $\dim S_k(\Gamma(N, M))$. The dimension of the space of cusp forms for $\Gamma(N, M)$ may also be computed via the Riemann-Roch theorem, an approach worked out in detail by Quer [27].

Another interesting application of the Eichler-Selberg trace formula is the ‘‘vertical’’ equidistribution of eigenvalues of Hecke operators acting on $S_k(\Gamma_0(N), \chi)$. It was proved independently around the same time by Conrey, Duke and Farmer [6] and Serre [32] that for p a fixed prime, as $k, N \rightarrow \infty$ through even weights k and levels N such that $p \nmid N$ that the eigenvalues of T_p become equidistributed with respect to the measure

$$d\mu_p = \frac{p+1}{\pi} \frac{(1-x^2/4)^{1/2}}{(p^{1/2} + p^{-1/2})^2 - x^2} dx$$

on $[-2, 2]$ (which is not the Sato-Tate measure). Our Theorem 9 should yield a similar equidistribution result for Hecke eigenvalues of T_p acting on the spaces $S_k(\Gamma(N, M))$.

We mention one more amusing application of the explicit formulas we prove, in particular an application of the simple case of Example 1. Let ℓ, p be two primes with $\ell > (\sqrt{p} + 1)^2$, which implies that $\ell \neq p$ and $p \not\equiv \pm 1 \pmod{\ell}$. For ℓ and p in this range, the Hasse bound

implies that an elliptic curve over \mathbb{F}_p cannot have a point of order ℓ and so $\mathcal{C}(\mathbb{Z}/\ell\mathbb{Z})$ is empty. Then for any two primes ℓ, p satisfying $\ell > (\sqrt{p} + 1)^2$, the result of Example 1 implies that

$$\mathrm{Tr}(T_p|S_2(\Gamma_1(\ell))) = p + 1 - \frac{\ell - 1}{2}.$$

Another interesting choice is to take ℓ and p to be primes such that $(\sqrt{p} - 1)^2 < \ell < (\sqrt{p} + 1)^2$ and $p - \ell \neq -1, 0, 1$. In this case we get a formula for a Hurwitz-Kronecker class number (for the definition of $H(\Delta)$ see (28)):

$$H((p + 1 - \ell)^2 - 4p) = -\frac{\mathrm{Tr}(T_p|S_2(\Gamma_1(\ell)))}{\ell - 1} + \frac{p + 1}{\ell - 1} - \frac{1}{2}.$$

These types of examples show that one should only expect equidistribution with respect to the Sato-Tate measure when $p \rightarrow \infty$ much faster than the conductor. It would be interesting to study with what uniformity one can expect Sato-Tate equidistribution when the conductor and p tend to infinity simultaneously.

2.5. Proofs of Theorems 4 and 5. We begin by using Theorem 6 to prove Theorems 4 and 5. Throughout this section we suppose that \mathbb{F}_q is a finite field of characteristic p , that $q = p^u$ for some integer $u \geq 1$.

Proof of Theorem 4. We split up the sum defining $\mathbb{E}_q(n_1(E))$ based on the value of $n_2(E)$:

$$(13) \quad \mathbb{E}_q(n_1(E)) = \sum_{1 \leq m \leq \sqrt{q}+1} \left(\frac{q+1}{m} \mathbb{E}_q(\Phi(n_2 = m)) - \frac{1}{m} \mathbb{E}_q(\Phi(n_2 = m)t_E) \right).$$

If $m \mid q - 1$ then applying Theorem 6 with $k = 3$ we have $\mathbb{E}_q(\Phi(n_2 = m)t_E) \ll_\varepsilon q^\varepsilon$. Applying Theorem 6 with $k = 2$ gives

$$(14) \quad \mathbb{E}_q(\Phi(n_2 = m)) = \frac{1}{\psi(m^2)\varphi(m)} \prod_{\substack{v \text{ prime} \\ v \mid \frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v(v^2 - 1)} \right) \prod_{\substack{v \text{ prime} \\ v \mid \frac{q-1}{m} \\ v \mid m}} \left(1 - \frac{1}{v^3} \right) + O_\varepsilon \left(q^{-\frac{1}{2} + \varepsilon} \right).$$

The error term here is uniform in m . If $m \nmid q - 1$ then $\mathcal{C}(n_2 = m)$ is empty and $\mathbb{E}_q(\Phi(n_2 = m)) = \mathbb{E}_q(\Phi(n_2 = m)t_E) = 0$.

We substitute (14) into (13) and also see that

$$\sum_{\substack{1 \leq m \leq \sqrt{q}+1 \\ m \mid q-1}} \frac{1}{m} \mathbb{E}_q(\Phi(n_2 = m)t_E) \ll_\varepsilon q^\varepsilon,$$

so it suffices to estimate

$$(15) \quad (q + 1) \sum_{\substack{1 \leq m \leq \sqrt{q}+1 \\ m \mid q-1}} \frac{\mathbb{E}_q(\Phi(n_2 = m))}{m}.$$

The error term from (14) makes a contribution of size $O_\varepsilon(q^{1/2+\varepsilon})$ to (15). The main term from (14) gives a main term of $(q+1)$ times

$$c(q) \stackrel{\text{def}}{=} \sum_{m|q-1} \frac{1}{m\psi(m^2)\varphi(m)} \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v(v^2-1)}\right) \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v^3}\right),$$

where we have extended the sum over m to all $m | q-1$ at a cost of a small error term. Note that $c(q)$ is a multiplicative function of $q-1$, so let $c(q) = f(q-1)$ where f is multiplicative. For a prime power ℓ^α we calculate

$$f(\ell^\alpha) = 1 - \frac{1}{\ell^2(\ell+1)} \left(1 + \frac{1}{\ell^4-1} - \frac{1}{\ell^{4\alpha-4}(\ell^4-1)}\right).$$

□

Proof of Theorem 5. We have

$$(16) \quad \mathbb{E}_q(n_2(E)) = \sum_{\substack{1 \leq m \leq \sqrt{q}+1 \\ m|q-1}} m \mathbb{E}_q(\Phi(n_2 = m)).$$

The calculation of $\mathbb{E}_q(\Phi(n_2 = m))$ in (14) gives a useful bound when $m \ll q^{1/6+\varepsilon}$. When m is larger, we need a trivial estimate. This estimate will involve bounds for Hurwitz-Kronecker class numbers $H(\Delta)$, which will be defined in Section 3.

For $m \geq 3$, we have by Lemma 2 in Section 3 below and Lemma 4.8 of [29] that

$$\left| q \mathbb{E}_q(\Phi(n_2 = m)) - \frac{p-1}{24} \delta(m, \sqrt{q}-1) - \frac{p-1}{24} \delta(m, \sqrt{q}+1) \right| \leq \sum_{\substack{t^2 < 4q \\ t \equiv q+1 \pmod{m^2}}} H\left(\frac{t^2 - 4q}{m^2}\right),$$

where the terms $\delta(m, \sqrt{q} \pm 1)$ are 0 if q is not a square, and the Hurwitz-Kronecker class number $H(\Delta)$ is defined in (28). To bound this expression we note that

$$H(\Delta) \ll |\Delta|^{1/2+\varepsilon}$$

by the class number formula and the upper bound $L(1, \chi) \ll \log |\Delta|$ for the associated L -functions. Then

$$(17) \quad \sum_{\substack{t^2 < 4q \\ t \equiv q+1 \pmod{m^2}}} H\left(\frac{t^2 - 4q}{m^2}\right) \ll_\varepsilon \frac{q^{1/2+\varepsilon}}{m} \left(1 + \frac{q^{1/2}}{m^2}\right).$$

We split (16) at $m = q^{1/6+\varepsilon}$ and use (14) when $m \leq q^{1/6+\varepsilon}$ and (17) when $m > q^{1/6+\varepsilon}$:

$$\begin{aligned} \mathbb{E}_q(n_2(E)) &= \sum_{m|q-1} \frac{m}{\psi(m^2)\varphi(m)} \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v(v^2-1)}\right) \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v^3}\right) + \frac{p}{12\sqrt{q}} \mathbf{1}_{q=\square} \\ &\quad + O_\varepsilon(q^{-1/3+\varepsilon}), \end{aligned}$$

where $\mathbb{1}_{q=\square} = 1$ if q is a square and is 0 otherwise. Setting

$$g(q-1) = \sum_{m|q-1} \frac{m}{\psi(m^2)\varphi(m)} \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v \nmid m}} \left(1 - \frac{1}{v(v^2-1)}\right) \prod_{\substack{v \text{ prime} \\ v|\frac{q-1}{m} \\ v|m}} \left(1 - \frac{1}{v^3}\right),$$

we see that $g(q-1) = b(q)$. It suffices to compute g on prime powers. We calculate

$$g(\ell^\alpha) = 1 + \frac{1}{\ell(\ell+1)} \frac{1 - \ell^{-2\alpha}}{1 - \ell^{-2}}.$$

□

2.6. Proof of Theorem 6. Let $A(md, md)$ denote the group $\mathbb{Z}/md\mathbb{Z} \times \mathbb{Z}/md\mathbb{Z}$. Note that $\mathcal{C}(A(md, md))$ is empty if $(md, q) > 1$ since the set of isomorphism classes over $\overline{\mathbb{F}}_q$, $\{E/\overline{\mathbb{F}}_q : A(md, md) \hookrightarrow E(\overline{\mathbb{F}}_q)\}$ is empty. Also, $\mathcal{C}(A(md, md))$ is empty if $md \nmid q-1$ by the first paragraph of the proof of Proposition 1, or by a Weil pairing argument.

In this proof we write $U(t, q) = U_{k-2}(t, q)$. We have that

$$(18) \quad \mathbb{E}_q(U(t_E, q)\Phi(n_2 = m)) = \sum_{d \geq 1} \mu(d) \mathbb{E}_q(U(t_E, q)\Phi_{A(md, md)}).$$

We assume that $m | q-1$ for the remainder of this argument, since otherwise $\mathbb{E}_q(U(t_E, q)\Phi(n_2 = m)) = 0$. When $m | q-1$, the sum on the right hand side of (18) is finite.

Applying Theorem 3 to the right hand side of (18) shows that

$$(19) \quad \mathbb{E}_q(U(t_E, q)\Phi(n_2 = m)) = \frac{1}{q} \sum_{d|\frac{q-1}{m}} \mu(d) (T_{md, md}(q, 1) - p^{k-1}T_{md, md}(q/p^2, p)) \\ + q^{k/2-1} \frac{(p-1)(k-1)}{24q} \sum_{d|\frac{q-1}{m}} \mu(d) (\delta_{md}(q^{1/2}, 1) + (-1)^k \delta_{md}(q^{1/2}, -1)).$$

By Möbius inversion, the second line of (19) is equal to

$$(20) \quad q^{k/2-1} \frac{(p-1)(k-1)}{24q} (\delta(q, (m+1)^2) + (-1)^k \delta(q, (m-1)^2)) \leq \frac{(k-1)}{12} q^{\frac{k-3}{2}}.$$

From the first line of (19) and the definitions in the introduction we have that

$$\frac{1}{q} \sum_{d|\frac{q-1}{m}} \mu(d) T_{md, md}(q, 1) = D - T + I - H,$$

where

$$D = \frac{1}{q} \sum_{d|\frac{q-1}{m}} \mu(d) \frac{\sigma(q)\delta(k, 2)}{\psi((md)^2)\varphi(md)},$$

$$T = \frac{1}{q} \sum_{d|\frac{q-1}{m}} \mu(d) \frac{\text{Tr}(T_q | S_k(\Gamma(md)))}{\psi((md)^2)\varphi(md)},$$

$$I = \frac{1}{q} \sum_{d|\frac{q-1}{m}} \mu(d) \frac{(k-1)}{24} q^{k/2-1} (\delta_{md}(q^{1/2}, 1) + (-1)^k \delta_{md}(q^{1/2}, -1)),$$

and

$$H = \frac{1}{4q} \sum_{d|\frac{q-1}{m}} \frac{\mu(d)}{\psi((md)^2)} \sum_{i=0}^v \min(p^i, p^{v-i})^{k-1} \sum'_{\substack{\tau|(md)^2 \\ g|p^i-p^{v-i}}} \varphi(g) (\delta_{md}(y_i, 1) + (-1)^k \delta_{md}(y_i, -1)),$$

where the ' on the sum means that $g = (\tau, (md)^2/\tau)$, and y_i is the unique element of $(\mathbb{Z}/(m^2d^2/g)\mathbb{Z})^\times$ such that $y_i \equiv p^i \pmod{\tau}$ and $y_i \equiv p^{v-i} \pmod{m^2d^2/(g\tau)}$.

We estimate each of the terms D, T, I , and H . A short calculation shows that

$$(21) \quad D = \frac{\sigma(q)}{q} \frac{1}{\psi(m^2)\varphi(m)} \prod_{\substack{\ell|\frac{q-1}{m} \\ \ell \nmid m}} \left(1 - \frac{1}{\ell(\ell^2-1)}\right) \prod_{\substack{\ell|\frac{q-1}{m} \\ \ell \nmid m}} \left(1 - \frac{1}{\ell^3}\right) \delta(k, 2).$$

We apply Deligne's bound on Hecke eigenvalues to T to get

$$|T| \leq 2q^{\frac{k-3}{2}} \log q \sum_{d|\frac{q-1}{m}} \frac{\dim S_k(\Gamma(md))}{\psi((md)^2)\varphi(md)}.$$

We have (e.g. [9, §3.9]) that

$$(22) \quad \dim S_k(\Gamma(N)) \leq \frac{kN^3}{12}.$$

Applying this bound we find

$$(23) \quad |T| \leq \frac{\zeta(2)}{6} kq^{\frac{k-3}{2}} \tau(q-1) (\log q).$$

Möbius inversion shows that

$$(24) \quad |I| = \frac{q^{\frac{k-4}{2}}(k-1)}{24} |\delta(q, (m+1)^2) + (-1)^k \delta(q, (m-1)^2)| \leq \frac{(k-1)}{12} q^{\frac{k-4}{2}}.$$

Lastly, we have

$$(25) \quad |H| \leq \frac{1}{q} \sum_{i=0}^v \min(p^i, p^{v-i})^{k-1} \sum_{d|\frac{q-1}{m}} \frac{1}{\psi((md)^2)} \sum_{\tau|md} \varphi(\tau) \leq 2q^{\frac{k-3}{2}} \tau(q-1) \log q.$$

Very similar estimates hold for the term

$$\frac{p^{k-1}}{q} \sum_{d|\frac{q-1}{m}} \mu(d) T_{md,md}(q/p^2, p)$$

of (19) so we omit the calculation. Drawing together (19), (20), (21), (23), (24), (25) we conclude the expression in the statement of Theorem 6.

3. COUNTING CURVES CONTAINING A PRESCRIBED SUBGROUP

Recall the definitions of the probability measure \mathbb{P}_q and of the sets \mathcal{C} , $\mathcal{C}(A)$, and $\mathcal{C}(A, t)$, from the introduction. In Section 1.3 we explained that our main goal is to give a formula for

$$(26) \quad \mathbb{E}_q(U_{k-2}(t_E, q)\Phi_A) = \frac{1}{q} \sum_{\substack{E/\mathbb{F}_q \\ A \hookrightarrow E(\mathbb{F}_q)}} \frac{U_{k-2}(t_E, q)}{\#\text{Aut}_{\mathbb{F}_q}(E)} = \sum_{t^2 \leq 4q} U_{k-2}(t, q)\mathbb{P}_q(\mathcal{C}(A, t)),$$

where $U_{k-2}(t, q)$ are the normalized Chebyshev polynomials defined in (1).

In this section we give formulas for $\mathbb{P}_q(\mathcal{C}(A, t))$ in terms of class numbers of orders in imaginary quadratic fields. In the special case that $A \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ such results are due to Deuring [8], Lenstra [25], Schoof [29], and Waterhouse [35]. The following is a weighted version of Theorem 4.6 of [29]. We begin with some definitions.

For $d < 0$ with $d \equiv 0, 1 \pmod{4}$, let $h(d)$ denote the class number of the unique quadratic order of discriminant d . Let

$$(27) \quad h_w(d) \stackrel{\text{def}}{=} \begin{cases} h(d)/3, & \text{if } d = -3, \\ h(d)/2, & \text{if } d = -4, \\ h(d) & \text{if } d < 0, d \equiv 0, 1 \pmod{4}, \text{ and } d \neq -3, -4 \\ 0 & \text{otherwise} \end{cases}$$

and for $\Delta \equiv 0, 1 \pmod{4}$ let

$$(28) \quad H(\Delta) \stackrel{\text{def}}{=} \sum_{d^2 | \Delta} h_w\left(\frac{\Delta}{d^2}\right)$$

be the Hurwitz-Kronecker class number. For $a \in \mathbb{Z}$ and n a positive integer, the Kronecker symbol $\left(\frac{a}{n}\right)$ is defined to be the completely multiplicative function in n such that if p is an odd prime $\left(\frac{a}{p}\right)$ is the quadratic residue symbol and if $p = 2$

$$(29) \quad \left(\frac{a}{2}\right) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } 2 \mid a, \\ 1 & \text{if } a \equiv \pm 1 \pmod{8}, \\ -1 & \text{if } a \equiv \pm 5 \pmod{8}. \end{cases}$$

Lemma 2. *Let $t \in \mathbb{Z}$. Suppose $q = p^v$ where p is prime and $v \geq 1$. Then if q is not a square*

$$\begin{aligned} \mathbb{P}_q(\mathcal{C}(1, t)) &= \frac{1}{2q} H(t^2 - 4q) && \text{if } t^2 < 4q \text{ and } p \nmid t, \\ &= \frac{1}{2q} H(-4p) && \text{if } t = 0, \\ &= \frac{1}{4q} && \text{if } t^2 = 2q \text{ and } p = 2, \\ &= \frac{1}{6q} && \text{if } t^2 = 3q \text{ and } p = 3, \end{aligned}$$

and if q is a square

$$\begin{aligned}\mathbb{P}_q(\mathcal{C}(1, t)) &= \frac{1}{2q} H(t^2 - 4q) && \text{if } t^2 < 4q \text{ and } p \nmid t, \\ &= \frac{1}{4q} \left(1 - \left(\frac{-4}{p} \right) \right) && \text{if } t = 0, \\ &= \frac{1}{6q} \left(1 - \left(\frac{-3}{p} \right) \right) && \text{if } t^2 = q, \\ &= \frac{p-1}{24q} && \text{if } t^2 = 4q,\end{aligned}$$

and $\mathbb{P}_q(\mathcal{C}(1, t)) = 0$ in all other cases.

Next we state separately the case $n = 2$. Lemma 3 is essentially Lemma 4.8 of [29]. We write $A_{2,2} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Lemma 3. *Let $q = p^v$ where $p \neq 2$ is prime. Suppose that $t \in \mathbb{Z}$ satisfies $t^2 \leq 4q$. If $q \equiv 3 \pmod{4}$ then*

$$\begin{aligned}\mathbb{P}_q(\mathcal{C}(A_{2,2}, t)) &= \frac{1}{2q} H\left(\frac{t^2 - 4q}{4}\right) && \text{if } p \nmid t \text{ and } t \equiv q + 1 \pmod{4}, \\ &= \frac{h_w(-p)}{2q} && \text{if } t = 0, \\ &= \mathbb{P}_q(\mathcal{C}(1, t)) && \text{if } t^2 = 4q,\end{aligned}$$

and if $q \equiv 1 \pmod{4}$

$$\begin{aligned}\mathbb{P}_q(\mathcal{C}(A_{2,2}, t)) &= \frac{1}{2q} H\left(\frac{t^2 - 4q}{4}\right) && \text{if } p \nmid t \text{ and } t \equiv q + 1 \pmod{4}, \\ &= \mathbb{P}_q(\mathcal{C}(1, t)) && \text{if } t^2 = 4q,\end{aligned}$$

and $\mathbb{P}_q(\mathcal{C}(1, t)) = 0$ in all other cases.

We also highlight the case where $A = A_{n,n} = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ with $n > 2$. This is a weighted version of Theorem 4.9 of [29].

Lemma 4. *Let $q = p^v$ where p is prime and $n > 2$ be a positive integer with $p \nmid n$. Suppose that $t \in \mathbb{Z}$ satisfies $t^2 \leq 4q$. If q is not a square then*

$$\begin{aligned}\mathbb{P}_q(\mathcal{C}(A_{n,n}, t)) &= \frac{1}{2q} H\left(\frac{t^2 - 4q}{n^2}\right) \delta_n(q, 1) && \text{if } p \nmid t, \text{ and } t \equiv q + 1 \pmod{n^2}, \\ &= 0 && \text{otherwise,}\end{aligned}$$

and if q is a square then

$$\begin{aligned}\mathbb{P}_q(\mathcal{C}(A_{n,n}, t)) &= \frac{1}{2q} H\left(\frac{t^2 - 4q}{n^2}\right) \delta_n(q, 1) && \text{if } p \nmid t, \text{ and } t \equiv q + 1 \pmod{n^2}, \\ &= \mathbb{P}_q(\mathcal{C}(1, 2\sqrt{q})) \delta_n(\sqrt{q}, 1) && \text{if } t = 2\sqrt{q}, \\ &= \mathbb{P}_q(\mathcal{C}(1, 2\sqrt{q})) \delta_n(\sqrt{q}, -1) && \text{if } t = -2\sqrt{q}, \\ &= 0 && \text{otherwise,}\end{aligned}$$

and $\mathbb{P}_q(\mathcal{C}(1, t)) = 0$ in all other cases.

We use an inclusion-exclusion argument to express $\mathcal{C}(A, t)$ in terms of the sets $\mathcal{C}(A_{n,n}, t)$ and congruence conditions on t and q . This extends Lemmas 2, 3, and 4 to a general finite abelian group A of rank at most 2. See Theorems 7 and 8.

We introduce a function $D(t; n)$ that plays a large role throughout the rest of the paper. Fix n_1 and n_2 with $n_2 \mid n_1$ and $d \in (\mathbb{Z}/n_1\mathbb{Z})^\times$. If $n' \mid n_1$ and $d^2q \equiv 1 \pmod{n'}$ then by Lemma 5 below $dq + d^{-1}$ is a well-defined residue class modulo n_1n' . For any $n \mid n_1n'$ we let

$$(30) \quad D(t; n) \stackrel{\text{def}}{=} \delta_n(dq + d^{-1}, t) = \begin{cases} 1 & \text{if } n \mid dq + d^{-1} - t \\ 0 & \text{otherwise.} \end{cases}$$

We must check that $D(t; n)$ is well-defined (à priori $dq + d^{-1}$ only makes sense modulo n_1), which we will do in Lemma 5. This lemma will be used extensively throughout Section 4 with various choices of parameters.

For $m \leq n$ and $c \in \mathbb{Z}/\ell^m\mathbb{Z}$ we call the pre-image of c in $\mathbb{Z}/\ell^n\mathbb{Z}$ under the canonical projection $\mathbb{Z}/\ell^n\mathbb{Z} \rightarrow \mathbb{Z}/\ell^m\mathbb{Z}$ the set of *lifts* modulo ℓ^n . Choosing a particular c_0 modulo ℓ^n that is a lift of c , we can describe the set of all lifts of c by $c_0 + j\ell^m \pmod{\ell^n}$ where $0 \leq j < \ell^{n-m}$.

Lemma 5. *Let $0 \leq B \leq C$, $D \in (\mathbb{Z}/\ell^C\mathbb{Z})^\times$, and $D^2q \equiv 1 \pmod{\ell^B}$. For any i satisfying $C \leq i \leq C + B$ we have that $Dq + D^{-1} \pmod{\ell^i}$ is the same residue class for any lift of D to a residue class modulo ℓ^i , and we say it is “well-defined”.*

Proof. We write the set of lifts of $D \pmod{\ell^C}$ to residue classes modulo ℓ^i as $D + D'\ell^C$ with $0 \leq D' < \ell^{i-C}$. Then

$$\begin{aligned} (D + D'\ell^C)q + (D + D'\ell^C)^{-1} &\equiv Dq + D^{-1} + (q - D^{-2})D'\ell^C \\ &\equiv Dq + D^{-1} \pmod{\ell^i} \end{aligned}$$

since $C \leq i \leq C + B \leq 2C$, completing the proof. \square

When $d = 1$, the condition that $D(t; n) = 0$ is the same as $n \nmid q + 1 - t$, in which case $\mathcal{C}(\mathbb{Z}/n\mathbb{Z}, t)$ is empty. Note that $D(t; n)$ is multiplicative for fixed t , i.e. if $(n, m) = 1$ then $D(t; n)D(t; m) = D(t; nm)$.

We say that $\nu \in \mathbb{N}$ is a full divisor of n_1 and write $\nu \parallel n_1$ if for all primes $\ell \mid \nu$ we have $v_\ell(\nu) = v_\ell(n_1)$. For $\mu, \nu \in \mathbb{N}$ we write $\mu \prec \nu$ if:

- (1) The integer μ is divisible by all the primes dividing ν and no others, i.e. $\mu \mid \nu^\infty$ and $\nu \mid \mu^\infty$, and
- (2) For all primes $\ell \mid \nu$ we have $v_\ell(\mu) \leq v_\ell(n_1/n_2) - 1$.

Finally, we define the function $D_{\nu,\mu}(t)$ (which also depends on q, d, n_1, n_2) to be

$$(31) \quad D_{\nu,\mu}(t) \stackrel{\text{def}}{=} \prod_{\ell \mid \nu} (D(t; \ell^{v_\ell(n_1n_2\mu)-1}) - D(t; \ell^{v_\ell(n_1n_2\mu)})).$$

When $d = 1$, $D_{\nu,\mu}(t) = 1$ if for each prime ℓ dividing μ , $v_\ell(q + 1 - t) = v_\ell(n_1n_2\mu) - 1$.

Now we define the following class numbers:

$$(32) \quad H_{n_1, n_2}(t, q, d) \stackrel{\text{def}}{=} \frac{1}{2} H\left(\frac{t^2 - 4q}{n_2^2}\right) \delta_{n_2}(d^2 q, 1) D(t; n_1 n_2) \\ + \sum_{\substack{m|n_1 \\ m \geq 2}} \sum_{\mu \prec m} \lambda(m) \frac{1}{2} H\left(\frac{t^2 - 4q}{(n_2 \mu)^2}\right) \delta_{n_2 \mu}(d^2 q, 1) D_{n_1, \mu}(t),$$

where $\lambda(m) = (-1)^{\omega(m)}$ and $\omega(m)$ denotes the number of distinct prime factors of m (so $\lambda(m)$ is almost the Liouville function).

Note in particular that when $n_1 = \ell^e$, and $n_2 = \ell^\delta$ we have

$$(33) \quad H_{\ell^e, \ell^\delta}(t, q, d) = \frac{1}{2} H\left(\frac{t^2 - 4q}{\ell^{2\delta}}\right) \delta_{\ell^\delta}(d^2 q, 1) D(t; \ell^{e+\delta}) \\ - \sum_{k=1}^{e-\delta-1} \frac{1}{2} H\left(\frac{t^2 - 4q}{\ell^{2(\delta+k)}}\right) \delta_{\ell^{\delta+k}}(d^2 q, 1) (D(t; \ell^{e+\delta+k-1}) - D(t; \ell^{e+\delta+k})).$$

We can divide the set of isomorphism classes of elliptic curve into those that are ordinary and those that are supersingular. We have

$$\mathbb{P}_q(\mathcal{C}(A, t)) = \begin{cases} \mathbb{P}_q(\mathcal{C}(A, t), E \text{ ordinary}) & \text{if } p \nmid t \\ \mathbb{P}_q(\mathcal{C}(A, t), E \text{ supersingular}) & \text{if } p \mid t, \end{cases}$$

and we deal with these two cases separately.

We now state one of the main results of this section, a formula for the number of isomorphism classes E/\mathbb{F}_q such that $E(\mathbb{F}_q)$ contains a subgroup isomorphic to A , has $t_E = t$, and is ordinary. We return to the supersingular case at the end of this section.

Theorem 7. *For a finite abelian group A of rank at most 2 we denote by $n_1(A)$ and $n_2(A)$ the first and second invariant factors of A , respectively. We have that*

$$\mathbb{P}_q(\mathcal{C}(A, t), E \text{ ordinary}) = \begin{cases} \frac{1}{q} H_{n_1(A), n_2(A)}(t, q, 1) & \text{if } p \nmid t \text{ and } t^2 < 4q \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mathbb{P}_q(\mathcal{C}(A, t)) = 0$ unless $q \equiv 1 \pmod{n_2(A)}$ by the definitions of these class numbers. The crux of the proof of Theorem 7 is the following proposition.

Proposition 2. *Let $A_{n_1, n_2} = \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$. We have*

$$(34) \quad \mathbb{P}_q(\mathcal{C}(A_{n_1, n_2}, t)) = \mathbb{P}_q(\mathcal{C}(A_{n_2, n_2}, t)) D(t; n_1 n_2) + \sum_{\substack{d|n_1 \\ d \geq 2}} \sum_{e \prec d} \lambda(d) \mathbb{P}_q(\mathcal{C}(A_{n_2 e, n_2 e}, t)) D_{n_1, e}(t),$$

where $\lambda(n) = (-1)^{\omega(n)}$ and $d = 1$ in the definition of $D(t; n)$.

Proof of Proposition 2. We prove the proposition by induction on the number of prime factors ℓ of n_1 for which $v_\ell(n_1) \neq v_\ell(n_2)$.

First consider the base case in which the number of such prime factors is 0, so $n_2 = n_1$. Then (34) holds since the sum over $e \prec d$ on the right hand side is empty, and the factor $D(t; n_1 n_2)$ is redundant since Lemmas 3 and 4 imply that $\mathcal{C}(A_{n_1, n_2}, t)$ is empty unless $n_1 n_2 \mid q + 1 - t$.

Let A be a finite abelian group of rank at most two. We suppose that (34) holds when n_1 has at most $\omega(|A|)$ prime factors \mathfrak{p} for which $v_{\mathfrak{p}}(n_1) \neq v_{\mathfrak{p}}(n_2)$, and show it also holds for n_1 having at most $\omega(|A|) + 1$ such prime factors. Let ℓ be a prime such that $\ell \nmid |A|$. For the rest of this proof we write $A_{e,d} = \mathbb{Z}/\ell^e\mathbb{Z} \times \mathbb{Z}/\ell^d\mathbb{Z}$ where $0 \leq d \leq e$. We are interested in the probability of the set of elliptic curves with $A \times A_{e,d} \hookrightarrow E(\mathbb{F}_q)$ and $t_E = t$. We note that $A \times A_{e,d} \hookrightarrow E(\mathbb{F}_q)$ if and only if $A \hookrightarrow E(\mathbb{F}_q)$ and $A_{e,d} \hookrightarrow E(\mathbb{F}_q)$.

Let

$$X \stackrel{\text{def}}{=} \mathcal{C}(A \times A_{e,d}) = \{E/\mathbb{F}_q : A \times A_{e,d} \hookrightarrow E(\mathbb{F}_q)\}.$$

We define a set of isomorphism classes that contains X :

$$X_0 \stackrel{\text{def}}{=} \{E/\mathbb{F}_q : A \times A_{d,d} \hookrightarrow E(\mathbb{F}_q) \text{ and } v_{\ell}(\#E(\mathbb{F}_q)) \geq d + e\}.$$

Let X_1 denote the difference of these sets, i.e. $X = X_0 \setminus X_1$.

If $E \in X_0$ and $v_{\ell}(\#E(\mathbb{F}_q)) \geq 2e - 1$ then $A_{e,d} \hookrightarrow E(\mathbb{F}_q)$, which implies $E \in \mathcal{C}(A \times A_{e,d})$. Therefore, each $E \in X_1$ satisfies $v_{\ell}(\#E(\mathbb{F}_q)) = d + e + k - 1$ for some k satisfying $1 \leq k \leq e - d - 1$. Let X_1^k denote the subset of X_1 with $v_{\ell}(\#E(\mathbb{F}_q)) = d + e + k - 1$, so that X_1 is a disjoint union of these sets.

If $E \in X_1^k$, then since $A_{e,d} \not\hookrightarrow E(\mathbb{F}_q)$ we have $A_{d+k,d+k} \hookrightarrow E(\mathbb{F}_q)$. Conversely, if $E \in X_0$ satisfies $A_{d+k,d+k} \hookrightarrow E(\mathbb{F}_q)$ and $v_{\ell}(\#E(\mathbb{F}_q)) = d + e + k - 1$, then $A_{e,d} \not\hookrightarrow E(\mathbb{F}_q)$ and $E \in X_1^k$. We conclude that

$$(35) \quad X_1^k = \{E/\mathbb{F}_q : A_{d+k,d+k} \hookrightarrow E(\mathbb{F}_q) \text{ and } v_{\ell}(\#E(\mathbb{F}_q)) = e + d + k - 1\}.$$

Fixing the value of t fixes $v_{\ell}(\#E(\mathbb{F}_q)) = v_{\ell}(q + 1 - t)$, so (35) implies that

$$\begin{aligned} \mathbb{P}_q(\mathcal{C}(A \times A_{e,d}, t)) &= \mathbb{P}_q(X_0, t_E = t) - \mathbb{P}_q(X_1, t_E = t) \\ &= \mathbb{P}_q(X_0, t_E = t) - \sum_{k=1}^{e-d-1} \mathbb{P}_q(X_1^k, t_E = t). \end{aligned}$$

Since $v_{\ell}(\#E(\mathbb{F}_q)) \geq d + e$ if and only if $D(t; \ell^{d+e}) = 1$, we have

$$\begin{aligned} &\mathbb{P}_q(X_0, t_E = t) - \sum_{k=1}^{e-d-1} \mathbb{P}_q(X_1^k, t_E = t) \\ &= \mathbb{P}_q(\mathcal{C}(A \times A_{d,d}, t))D(t; \ell^{d+e}) - \sum_{\substack{1 \leq k \leq e-d-1 \\ v_{\ell}(q+1-t) = e+d+k-1}} \mathbb{P}_q(\mathcal{C}(A \times A_{d+k,d+k}, t)). \end{aligned}$$

The induction hypothesis is now applicable to each probability on the right hand side of this expression, from which we deduce the proposition for n_1 having $\omega(|A|) + 1$ prime factors for which $v_{\mathfrak{p}}(n_1) \neq v_{\mathfrak{p}}(n_2)$. \square

Proof of Theorem 7. We apply Lemmas 2, 3, and 4 to (34). Rewriting this expression using definition of $H_{n_1, n_2}(t, q, d)$ completes the proof. \square

In order to prove a result analogous to Theorem 7 for supersingular curves, we define a version of the class numbers $H_{n_1, n_2}(t, q, d)$ that collects all of the contributions from supersingular curves. This involves an analysis of many special cases.

We define a function $H_{n_1, n_2}^*(t, q, d)$ as follows. If $n_2 > 2$ we define

$$(36) \quad H_{n_1, n_2}^*(t, q, d) \stackrel{\text{def}}{=} 0,$$

and if $n_2 = 2$ we define

$$(37) \quad H_{n_1, 2}^*(t, q, d) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{2} h_w(-p) D(t; 2n_1) & \text{if } t = 0 \\ 0 & \text{otherwise.} \end{cases}$$

If $n_2 = 1$ and $q = p^v$ with v even we define

$$(38) \quad H_{n_1, 1}^*(t, q, d) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{4} \left(1 - \left(\frac{-4}{p} \right) \right) D(t; n_1) & \text{if } t = 0, \\ \frac{1}{6} \left(1 - \left(\frac{-3}{p} \right) \right) D(t; n_1) & \text{if } t^2 = q, \\ 0 & \text{otherwise.} \end{cases}$$

If $n_2 = 1$ and $q = p^v$ with v odd then we define

$$(39) \quad H_{n_1, 1}^*(t, q, d) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{2} H(-4p) D(t; n_1) - \delta_4(n_1, 0) \frac{1}{2} h_w(-p) (D(t; n_1) - D(t; 2n_1)) & \text{if } t = 0, \\ \frac{1}{4} D(t; n_1) & \text{if } t^2 = 2q \text{ and } p = 2, \\ \frac{1}{6} D(t; n_1) & \text{if } t^2 = 3q \text{ and } p = 3, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 8. *For a finite abelian group A of rank at most 2 we denote by $n_1(A)$ and $n_2(A)$ the first and second invariant factors of A , respectively. We have that*

$$\mathbb{P}_q(\mathcal{C}(A, t), E \text{ supersingular}) = \begin{cases} \frac{1}{q} H_{n_1(A), n_2(A)}^*(t, q, 1) & \text{if } t^2 < 4q \\ \frac{p-1}{24q} \delta_{n_1(A)}(\sqrt{q}, \pm 1) & \text{if } t^2 = 4q. \end{cases}$$

Proof. We recall that E/\mathbb{F}_q is supersingular if and only if $p \mid t_E$. Lemma 2 implies that such a curve must have $t_E^2 \in \{0, q, 2q, 3q, 4q\}$, where $t_E^2 = 2q$ arises only when $p = 2$ and $t_E^2 = 3q$ arises only when $p = 3$. Lemmas 3 and 4 imply that there are not many different group structures to consider. We divide our argument into cases based on the value of n_2 .

The case $t^2 < 4q$ and $n_2(A) > 2$.

Lemma 4 implies that for any $e > 2$, $\mathbb{P}_q(\mathcal{C}(A_{e,e}, t), E \text{ supersingular}) = 0$. This implies that $\mathbb{P}_q(\mathcal{C}(A, t), E \text{ supersingular}) = 0$, which matches the definition of $H_{n_1, n_2}^*(t, q, 1)$ given by (36).

The case $t^2 < 4q$ and $n_2(A) = 2$.

Lemma 3 implies that in this case $\mathbb{P}_q(\mathcal{C}(A_{2,2}, t), E \text{ supersingular}) = 0$ unless $t = 0$ and $q \equiv 3 \pmod{4}$. Lemma 4 implies that for any $e \geq 2$, $\mathbb{P}_q(\mathcal{C}(A_{2e, 2e}, t), E \text{ supersingular}) = 0$. Therefore,

$$\mathbb{P}_q(\mathcal{C}(A_{n_1, 2}, t), E \text{ supersingular}) = \mathbb{P}_q(\mathcal{C}(A_{2, 2}, t), E \text{ supersingular}).$$

The expression in Lemma 3 matches the definition of $H_{n_1, 2}^*(t, q, 1)$ given by (37).

The case $t^2 < 4q$ and $n_2(A) = 1$.

Lemma 4 implies that for any $e \geq 3$, $\mathbb{P}_q(\mathcal{C}(A_{e,e}, t), E \text{ supersingular}) = 0$. As above, Lemma 3 implies that $\mathbb{P}_q(\mathcal{C}(A_{2,2}, t), E \text{ supersingular}) = 0$ unless $t = 0$ and $q \equiv 3 \pmod{4}$. If $2 \prec d$ then d is a power of 2 and $\lambda(d) = -1$. By Proposition 2 we have

$$\begin{aligned} & \mathbb{P}_q(\mathcal{C}(A_{n_1,1}, t), E \text{ supersingular}) \\ &= \mathbb{P}_q(\mathcal{C}(1, t), E \text{ supersingular}) D(t; n_1) - \delta_4(n_1, 0) \mathbb{P}_q(\mathcal{C}(A_{2,2}, t), E \text{ supersingular}) D_{n_1,2}(t). \end{aligned}$$

Applying Lemmas 2 and 3 to this expression we check that it matches the remaining cases in the definition of $H_{n_1,1}^*(t, q, 1)$.

The case $t^2 = 4q$.

In this case any E/\mathbb{F}_q with $t_E = t = \pm 2\sqrt{q}$ has

$$E(\mathbb{F}_q) \cong \mathbb{Z}/(\sqrt{q} \mp 1)\mathbb{Z} \times \mathbb{Z}/(\sqrt{q} \mp 1)\mathbb{Z}$$

by [29, Lemma 4.8(ii)]. Therefore

$$\mathbb{P}_q(\mathcal{C}(A_{n_1, n_2}, \pm 2\sqrt{q})) = \mathbb{P}_q(\mathcal{C}(1, \pm 2\sqrt{q})) \delta_{n_1}(\pm\sqrt{q}, 1).$$

Applying Lemma 2 to this expression concludes the proof of the theorem. \square

We summarize the main results of this section in the following proposition. We define two functions $\omega_A(q, d)$ and $\omega_A^*(q, d)$ that will be used extensively in Section 5. Let

$$(40) \quad \omega_A(q, d) \stackrel{\text{def}}{=} \sum_{\substack{t^2 < 4q \\ p \nmid t}} U_{k-2}(t, q) H_{n_1(A), n_2(A)}(t, q, d),$$

and

$$(41) \quad \omega_A^*(q, d) \stackrel{\text{def}}{=} \sum_{t^2 < 4q} U_{k-2}(t, q) H_{n_1(A), n_2(A)}^*(t, q, d).$$

Note also $U_{k-2}(\pm 1) = (k-1)(\pm 1)^k$. By (26) and Theorems 7 and 8 we have the following.

Proposition 3. *We have*

$$\begin{aligned} \mathbb{E}_q(U_{k-2}(t, q) \Phi_A) &= \frac{1}{q} \omega_A(q, d) + \frac{1}{q} \omega_A^*(q, 1) \\ &\quad + q^{k/2-1} \frac{(p-1)(k-1)}{24q} (\delta_{n_1(A)}(\sqrt{q}, 1) + (-1)^k \delta_{n_1(A)}(\sqrt{q}, -1)). \end{aligned}$$

Note that the right hand side above is 0 if $q \not\equiv 1 \pmod{n_2(A)}$. Indeed, the congruence $q \equiv 1 \pmod{n_2}$ is a necessary condition for $E(\mathbb{F}_q)$ to have a subgroup isomorphic to $\mathbb{Z}/n_2\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$ by, say, the Weil pairing on E .

4. TRACE FORMULAS

In this section we derive an Eichler-Selberg trace formula for the congruence subgroups $\Gamma(N, M)$, which were defined in equation (4). Our proof starts from the corresponding trace formula for $\Gamma_0(MN)$ with nebentype character. Let $S_k(\Gamma(N, M))$ be the space of weight $k \geq 2$ cusp forms for $\Gamma(N, M)$. (Here we follow Sections 5.1 and 5.2 of the book of Diamond and Shurman [9].) We have an exact sequence

$$1 \rightarrow \Gamma(N, M) \rightarrow \Gamma_0(NM) \rightarrow (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow 1.$$

The congruence subgroup $\Gamma_0(NM)$ acts on $S_k(\Gamma(N, M))$ via the slash operator with $\Gamma(N, M)$ acting trivially, so this action is via the quotient $(\mathbb{Z}/N\mathbb{Z})^\times$. Thus for each $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ we have the diamond operator

$$\langle d \rangle : S_k(\Gamma(N, M)) \rightarrow S_k(\Gamma(N, M))$$

given by

$$\langle d \rangle f = f|_\gamma, \text{ for any } \gamma = \begin{pmatrix} a & b \\ c & \delta \end{pmatrix} \in \Gamma_0(NM) \text{ with } \delta \equiv d \pmod{N}.$$

For a Dirichlet character $\chi \pmod{N}$ (considered as a imprimitive character \pmod{NM}) the space of cusp forms of $\Gamma_0(NM)$ with nebentype character χ is the χ -eigenspace of the diamond operators:

$$S_k(\Gamma_0(NM), \chi) = \{f \in S_k(\Gamma(N, M)) \text{ s.t. } \langle d \rangle f = \chi(d)f \text{ for all } d \in (\mathbb{Z}/N\mathbb{Z})^\times\}.$$

We also define Hecke operators for this group. For p prime, let

$$\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}.$$

Take a coset decomposition for the double coset

$$\Gamma(N, M)\alpha\Gamma(N, M) = \bigsqcup_j \Gamma(N, M)\beta_j,$$

where $\beta_j \in \mathrm{GL}_2^+(\mathbb{Q})$. Then

$$T_p : S_k(\Gamma(N, M)) \rightarrow S_k(\Gamma(N, M))$$

is defined by

$$T_p f = \sum_j f|_{\beta_j}.$$

So long as $(p^v, N) = 1$ we define the Hecke operators for prime powers inductively via

$$T_{p^v} = T_p T_{p^{v-1}} - p^{k-1} \langle p \rangle T_{p^{v-2}}.$$

One can check that the diamond operators and Hecke operators commute with each other and that T_p respects the decomposition

$$S_k(\Gamma(N, M)) = \bigoplus_{\chi \pmod{N}} S_k(\Gamma_0(NM), \chi).$$

Comparing Propositions 5.2.2 and 5.3.1 of [9] and Proposition 3.46 in the book of Knightly and Li [22] shows that the Hecke operators defined above coincide with those of [9]. Thus if $(dq, NM) = 1$ we have that

$$(42) \quad \text{Tr}(\langle d \rangle T_q | S_k(\Gamma(N, M))) = \sum_{\chi \pmod{NM}} \chi(d) \text{Tr}(T_q | S_k(\Gamma_0(NM), \chi)).$$

We use (42) to prove a trace formula for Hecke operators on the groups $\Gamma(N, M)$.

Theorem 9 (Eichler-Selberg Trace Formula for $\Gamma(N, M)$). *Let q, M, N be positive integers, $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, and $k \geq 2$ an integer. Let T_q be the q^{th} Hecke operator acting on $S_k(\Gamma(N, M))$ and $\langle d \rangle$ be the d -diamond operator acting on $S_k(\Gamma(N, M))$.*

Suppose that $M \mid N$, $(N, q) = 1$, and $d^2q \equiv 1 \pmod{M}$. Let $L = (d^2q - 1, N)$. We have that

$$\begin{aligned} \frac{\text{Tr}(\langle d \rangle T_q | S_k(\Gamma(N, M)))}{\varphi(N)} &= \frac{k-1}{24} q^{k/2-1} \psi(NM) (\delta_N(q^{1/2}d, 1) + (-1)^k \delta_N(q^{1/2}d, -1)) \\ &\quad - \frac{\psi(N^2)}{\psi(N^2/M^2)} \sum_{\Lambda|(L/M)} \frac{\varphi(\Lambda^2)\varphi(N/(M\Lambda))}{\varphi(N/M)} \sum_{t^2 < 4q} U_{k-2}(t, q) H_{N, \Lambda M}(t, q, d) \\ &\quad - \frac{1}{4} \sum_{b|q} \min(b, q/b)^{k-1} \sum_{\substack{\tau|NM \\ g|(b-q/b)}} \frac{\varphi(g)\varphi(N(M, g)/g)}{\varphi(N)} \\ &\quad \quad \times \left(\delta_{N(M, g)/g}(y_\tau d, 1) + (-1)^k \delta_{N(M, g)/g}(y_\tau d, -1) \right) \\ &\quad + \frac{\sigma(q)}{\varphi(N)} \delta(k, 2), \end{aligned}$$

where

- $U_{k-2}(t, q)$ is a normalized Chebyshev polynomial (see (1)),
- $g = (\tau, NM/\tau)$,
- y_τ is the unique element of $(\mathbb{Z}/(NM/g)\mathbb{Z})^\times$ such that $y_\tau \equiv b \pmod{\tau}$ and $y_\tau \equiv q/b \pmod{NM/\tau}$.

We will prove Theorem 9 using the trace formula for $\Gamma_0(N)$ with nebentypus. We refer to [22] for the proof and also to [21] for the statement.

Theorem 10 (Eichler-Selberg Trace Formula for $\Gamma_0(N), \chi$). *Let $k \geq 2$ and $\chi(-1) = (-1)^k$. We have*

$$(43) \quad \begin{aligned} \text{Tr}(T_q | S_k(\Gamma_0(N), \chi)) &= \frac{k-1}{12} \psi(N) \chi(q^{1/2}) q^{k/2-1} \\ &\quad - \frac{1}{2} \sum_{t^2 < 4q} U_{k-2}(t, q) \sum_{m^2 | (t^2 - 4q)} h_w \left(\frac{t^2 - 4q}{m^2} \right) \mu_\chi(t, m, q) \\ &\quad - \frac{1}{2} \sum_{b|q} \min(b, q/b)^{k-1} \sum_{\tau} \varphi((\tau, N/\tau)) \chi(y_\tau) \\ &\quad + \delta(k, 2) \mathbf{1}_{\chi=1} \sum_{\substack{c|q \\ (N, q/c)=1}} c \end{aligned}$$

where:

- $\chi(q^{1/2}) = 0$ if q is not a perfect square,
- $U_{k-2}(t, q)$ is a normalized Chebyshev polynomial (see (1)),
-

$$(44) \quad \mu_\chi(t, m, q) = \frac{\psi(N)}{\psi(N/(N, m))} \sum_{c \in (\mathbb{Z}/N\mathbb{Z})^\times}^{*m} \chi(c),$$

where the $*_m$ indicates that c runs through all elements of $(\mathbb{Z}/N\mathbb{Z})^\times$ that lift to solutions of $c^2 - tc + q \equiv 0 \pmod{N(N, m)}$,

- τ runs through all positive divisors of N such that $(\tau, N/\tau)$ divides both $N/\text{cond}(\chi)$ and $b - q/b$ where $\text{cond}(\chi)$ is the conductor of χ ,
- y_τ is the unique element of $\mathbb{Z}/(N/(\tau, N/\tau))\mathbb{Z}$ such that $y_\tau \equiv d \pmod{\tau}$ and $y_\tau \equiv q/d \pmod{N/\tau}$,
- and $\mathbf{1}_{\chi=1}$ is 1 if χ is the trivial character and is 0 otherwise.

Proof of Theorem 9. We define the following four terms, which correspond to the terms appearing on the right hand side of Theorem 10 applied to the group $\Gamma_0(MN)$:

(1)

$$T_\chi^{(i)} \stackrel{\text{def}}{=} \frac{k-1}{12} \psi(MN) \chi(q^{1/2}) q^{k/2-1},$$

(2)

$$T_\chi^{(e)} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{t^2 < 4q} U_{k-2}(t, q) T_\chi^{(e)}(t),$$

where

$$T_\chi^{(e)}(t) \stackrel{\text{def}}{=} \sum_{m^2 | (t^2 - 4q)} h_w \left(\frac{t^2 - 4q}{m^2} \right) \mu_\chi(t, m, q),$$

(3)

$$T_\chi^{(h)} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{b|q} \min(b, q/b)^{k-1} T_\chi^{(h)}(b),$$

where

$$T_\chi^{(h)}(b) \stackrel{\text{def}}{=} \sum_{\tau} \varphi((\tau, MN/\tau)) \chi(y),$$

(4)

$$T_\chi^{(d)} \stackrel{\text{def}}{=} \delta(k, 2) \mathbf{1}_{\chi=1} \sum_{\substack{c|q \\ (MN, q/c)=1}} c.$$

Then by (42), Theorem 10, and the fact that $S_k(\Gamma_0(MN), \chi)$ is $\{0\}$ unless $\chi(-1) = (-1)^k$ we have that

$$(45) \quad \frac{1}{\varphi(N)} \text{Tr}(\langle d \rangle T_q | S_k(\Gamma(N, M))) = T^{(i)} - T^{(e)} - T^{(h)} + T^{(d)}$$

where

(1)

$$T^{(i)} \stackrel{\text{def}}{=} \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) T_\chi^{(i)},$$

(2)

$$T^{(e)} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{t^2 < 4q} U_{k-2}(t, q) T^{(e)}(t),$$

where

$$T^{(e)}(t) \stackrel{\text{def}}{=} \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) T_\chi^{(e)}(t),$$

(3)

$$T^{(h)} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{b|q} \min(b, q/b)^{k-1} T^{(h)}(b),$$

where

$$T^{(h)}(b) \stackrel{\text{def}}{=} \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) T_\chi^{(h)}(b),$$

(4)

$$T^{(d)} \stackrel{\text{def}}{=} \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) T_\chi^{(d)}.$$

We compute each of the $T^{(*)}$ in a separate section and check that they match what is claimed in Theorem 9.

4.1. Identity Term. We have the orthogonality relation

$$(46) \quad \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) = \frac{1}{2} (\delta_N(d, 1) + (-1)^k \delta_N(d, -1)),$$

which shows that

$$(47) \quad \begin{aligned} T^{(i)} &= \frac{k-1}{12} q^{k/2-1} \psi(MN) \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(dq^{1/2}) \\ &= \frac{k-1}{24} q^{k/2-1} \psi(MN) (\delta_N(q^{1/2}d, 1) + (-1)^k \delta_N(q^{1/2}d, -1)). \end{aligned}$$

This matches the claimed identity term of Theorem 9.

4.2. Elliptic Term. The elliptic term $T^{(e)}$ is by far the most difficult to compute of the four terms in the trace formula. We start with the case that N is a power of a prime ℓ with $\ell \nmid q$. The goal of this section is to reduce the proof of this prime power case to the proof of Proposition 4, which we state below. We prove Proposition 4 in Section 4.5. Finally, in Section 4.6 we consider the case where N is not necessarily a prime power.

Let $0 \leq \alpha \leq \beta \leq \gamma$ be non-negative integers such that $(d^2q - 1, \ell^\gamma) = \ell^\beta$. Let $\Delta = t^2 - 4q$. Taking $N = \ell^\gamma$ and $M = \ell^\alpha$ in the definition of $T^{(e)}(t)$ we have that

$$(48) \quad T^{(e)}(t) = \sum_{m^2|\Delta} h_w \left(\frac{\Delta}{m^2} \right) \frac{1}{\varphi(\ell^\gamma)} \sum_{\substack{\chi \pmod{\ell^\gamma} \\ \chi(-1)=(-1)^k}} \chi(d) \mu_\chi(t, m, q).$$

Also let $\kappa = v_\ell(m)$ and $\nu = v_\ell(\Delta)$ so that $\kappa \leq \lfloor \nu/2 \rfloor$. We then have $(m, \ell^{\alpha+\gamma}) = \ell^{\min(\kappa, \alpha+\gamma)}$, which occurs often below. Let

$$(49) \quad W(d) \stackrel{\text{def}}{=} \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times} {}^*m \delta_{\ell^\gamma}(c, d^{-1})$$

where the *m notation was explained in the statement of Theorem 10. Note that $W(d)$ not only depends on d , but also on ℓ, γ , and α .

Using the orthogonality relation (46) we have

$$(50) \quad \begin{aligned} \frac{1}{\varphi(\ell^\gamma)} \sum_{\substack{\chi \pmod{\ell^\gamma} \\ \chi(-1)=(-1)^k}} \chi(d) \mu_\chi(t, m, q) &= \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma}/\ell^{\min(\kappa, \alpha+\gamma)})} \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times} {}^*m \frac{1}{\varphi(\ell^\gamma)} \sum_{\substack{\chi \pmod{\ell^\gamma} \\ \chi(-1)=(-1)^k}} \chi(dc) \\ &= \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma-\min(\kappa, \alpha+\gamma)})} \frac{1}{2} (W(d) + (-1)^k W(-d)). \end{aligned}$$

We also set

$$(51) \quad C(t, q, d) \stackrel{\text{def}}{=} \sum_{m^2|(t^2-4q)} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma-\min(\kappa, \alpha+\gamma)})} W(d)$$

so that by (48) and (50) we have

$$T^{(e)}(t) = \frac{1}{2} (C(t, q, d) + (-1)^k C(t, q, -d)).$$

Note that $C(t, q, d)$ not only depends on t, q, d but also on $\ell, \alpha, \beta, \gamma$.

Recall the definition of $H_{n_1, n_2}(t, q, d)$ from (33). The proof of the following proposition is the subject of Section 4.5.

Proposition 4. *We have*

$$C(t, q, d) = 2 \frac{\psi(\ell^{2\gamma})}{\psi(\ell^{2\gamma-2\alpha})} \sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j}) \varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} H_{\ell^\gamma, \ell^{\alpha+j}}(t, q, d).$$

Since $U_{k-2}(t, q)$ is an even (resp. odd) function of t when k is even (resp. odd) and $H_{n_1, n_2}(t, q, -d) = H_{n_1, n_2}(-t, q, d)$ we have

$$(52) \quad \sum_{t^2 < 4q} U_{k-2}(t, q) C(t, q, d) = (-1)^k \sum_{t^2 < 4q} U_{k-2}(t, q) C(t, q, -d).$$

By Proposition 4, (45), and the definitions that follow

$$T^{(e)} = \frac{\psi(\ell^{2\gamma})}{\psi(\ell^{2\gamma-2\alpha})} \sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} \sum_{t^2 < 4q} U_{k-2}(t, q) H_{\ell^\gamma, \ell^{\alpha+j}}(t, q, d),$$

matching exactly what is claimed in Theorem 9.

4.3. Hyperbolic Term. Let $\tau \mid MN$ and $g = (\tau, NM/\tau)$. We have that

$$T^{(h)}(b) = \frac{1}{\varphi(N)} \sum_{\substack{\chi \pmod{N} \\ \chi(-1) = (-1)^k}} \chi(d) \sum_{\substack{\tau \mid MN \\ g \mid (b-q/b) \\ g \mid (MN/\text{cond}(\chi))}} \varphi(g)\chi(y_\tau)$$

where y_τ is defined to be the unique element of $(\mathbb{Z}/(MN/g)\mathbb{Z})^\times$ that satisfies $y_\tau \equiv b \pmod{\tau}$ and $y_\tau \equiv q/b \pmod{MN/\tau}$. Swapping the order of summation we have

$$T^{(h)}(b) = \sum_{\substack{\tau \mid MN \\ g \mid (b-q/b)}} \frac{\varphi(g)}{\varphi(N)} \sum_{\substack{c \mid N \\ g \mid (MN/c)}} \sum_{\substack{\text{cond}(\chi) = c \\ \chi(-1) = (-1)^k}} \chi(dy_\tau).$$

Recall the following almost-orthogonality relation for sums of characters of a given conductor (see e.g. [11, Section 2])

$$(53) \quad \sum_{\substack{\text{cond}(\chi) = c \\ \chi(-1) = (-1)^k}} \chi(a) = \frac{1}{2} \sum_{\delta \mid (a-1, c)} \varphi(\delta) \mu\left(\frac{c}{\delta}\right) + (-1)^k \frac{1}{2} \sum_{\delta \mid (a+1, c)} \varphi(\delta) \mu\left(\frac{c}{\delta}\right).$$

Let

$$V(\tau, d) \stackrel{\text{def}}{=} \sum_{\substack{c \mid N \\ g \mid (MN/c)}} \sum_{\delta \mid (dy_\tau - 1, c)} \varphi(\delta) \mu\left(\frac{c}{\delta}\right)$$

so that we have

$$T^{(h)}(b) = \frac{1}{2} \sum_{\substack{\tau \mid MN \\ g \mid (b-q/b)}} \frac{\varphi(g)}{\varphi(N)} (V(\tau, d) + (-1)^k V(\tau, -d)).$$

By commutativity and associativity of Dirichlet convolution we have

$$(54) \quad V(\tau, d) = \sum_{\substack{a_1 \mid N \\ a_1 \mid dy_\tau - 1}} \varphi(a_1) \sum_{\substack{a_2 \mid (N/a_1) \\ g \mid a_2 M}} \mu\left(\frac{N}{a_1 a_2}\right).$$

For any positive integers n, g, M we have

$$(55) \quad \sum_{\substack{a \mid n \\ g \mid aM}} \mu\left(\frac{n}{a}\right) = \sum_{\substack{a \mid n \\ (g/(g, M)) \mid a}} \mu\left(\frac{n}{a}\right) = \delta_{(g/(g, M))}(n, 0) \sum_{a \mid (ng/(g, M))} \mu\left(\frac{ng}{a(g, M)}\right) = \delta\left(n, \frac{g}{(g, M)}\right),$$

and applying this to (54)

$$(56) \quad V(\tau, d) = \varphi \left(\frac{N(g, M)}{g} \right) \delta_{N(g, M)/g}(dy_\tau, 1).$$

Therefore we have

$$T^{(h)}(b) = \frac{1}{2} \sum_{\substack{\tau | MN \\ g | (b-q/b)}} \frac{\varphi(g)}{\varphi(N)} \varphi \left(\frac{N(g, M)}{g} \right) (\delta_{N(g, M)/g}(dy_\tau, 1) + (-1)^k \delta_{N(g, M)/g}(dy_\tau, -1)),$$

as claimed in the statement of Theorem 9.

4.4. Dual Term. We have immediately that

$$T^{(d)} = \frac{\sigma(q)}{\varphi(N)} \delta(k, 2)$$

as all terms in the sum over χ vanish except the identity character, and $(N, q) = 1$ by hypothesis. \square

4.5. The Proof of Proposition 4. In this section we compute $C(t, q, d)$, which is defined as a sum over divisors m^2 of Δ . We split up this sum based on the value of $v_\ell(m)$. Let

$$(57) \quad c_\kappa(t, q, d) \stackrel{\text{def}}{=} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma - \min(\kappa, \alpha+\gamma)})} W(d)$$

so that

$$(58) \quad C(t, q, d) = \sum_{\kappa \geq 0} c_\kappa(t, q, d).$$

Our goal is to rewrite these $c_\kappa(t, q, d)$ so that we can express $C(t, q, d)$ as a sum involving the class numbers $H_{\ell^\gamma, \ell^{\alpha+j}}(t, q)$ defined in Section 3. We start with some preliminary lemmas.

4.5.1. Preliminary Lemmas. We prove several lemmas that will be used in the proof of Proposition 4. Our first goal in simplifying $c_\kappa(t, q, d)$ is to compute $W(d)$. We do this in Lemma 9. Lemmas 6, 7, and 8 are technical and will be used in the proof of Lemma 9.

Lemma 6 allows us to convert the sums over solutions to the Hecke polynomial $c^2 - tc + q$ implicit in the $*_m$ notation in (49) into the indicator function of t lying in a certain residue class. We will apply this lemma for for several different choices of B, C , and D .

Lemma 6. *Let $0 \leq B \leq C \leq i \leq B + C$, $D \in (\mathbb{Z}/\ell^C \mathbb{Z})^\times$, $D^2 q \equiv 1 \pmod{\ell^B}$, and $t \in \mathbb{Z}$. The following statements are equivalent:*

- (1) *There exists a lift c modulo ℓ^i of $D^{-1} \pmod{\ell^C}$ that satisfies $c^2 - tc + q \equiv 0 \pmod{\ell^i}$.*
- (2) *The integer t satisfies $t \equiv Dq + D^{-1} \pmod{\ell^i}$, which is well-defined by Lemma 5.*

If the above equivalent statements hold, then every one of the ℓ^{i-C} lifts of $D^{-1} \pmod{\ell^C}$ to residue classes c modulo ℓ^i satisfies $c^2 - tc + q \equiv 0 \pmod{\ell^i}$.

Proof. Suppose that $c = D^{-1} + c_0 \ell^C$ is a solution to $c^2 - tc + q \equiv 0 \pmod{\ell^i}$ for some c_0 satisfying $0 \leq c_0 < \ell^{i-C}$. Then we solve for t :

$$\begin{aligned} t &\equiv (D^{-1} + c_0 \ell^C)^{-1} (q + (D^{-1} + c_0 \ell^C)^2) \pmod{\ell^i} \\ &\equiv Dq + D^{-1} \pmod{\ell^i}. \end{aligned}$$

Now assume that $t \equiv Dq + D^{-1} \pmod{\ell^i}$. We parametrize lifts of D^{-1} to $\mathbb{Z}/\ell^i\mathbb{Z}$ by $D^{-1} + c'\ell^C$ with $0 \leq c' < \ell^{i-C}$. Then

$$(D^{-1} + c'\ell^C)^2 - (Dq + D^{-1})(D^{-1} + c'\ell^C) + q \equiv (D^{-1} - Dq)c'\ell^C \equiv 0 \pmod{\ell^i}$$

for any value of c' . Thus, all ℓ^{i-C} lifts of $D^{-1} \pmod{\ell^C}$ are solutions to this quadratic polynomial modulo ℓ^i . \square

The second technical lemma, Lemma 7, will be used both in the proof of Lemma 9 and also repeatedly in the proof of Proposition 4. Recall $\nu = v_\ell(\Delta) = v_\ell(t^2 - 4q)$.

- Lemma 7.** (1) If $\nu < 2\beta$ then $t \not\equiv dq + d^{-1} \pmod{\ell^{\nu+1}}$.
(2) If $\ell = 2$ and $\nu < 2\beta$ then $t \not\equiv dq + d^{-1} \pmod{2^\nu}$.
(3) If $\beta < \gamma$ and $\nu \neq 2\beta$ then $t \not\equiv dq + d^{-1} \pmod{\ell^{2\beta+1}}$.

Proof. (1) If $\nu < 2\beta$ then $dq + d^{-1} \pmod{\ell^{\nu+1}}$ is well-defined by Lemma 5. If $t \equiv dq + d^{-1} \pmod{\ell^{\nu+1}}$ then $\Delta \equiv (dq - d^{-1})^2 \pmod{\ell^{\nu+1}}$. By definition $(d^2q - 1, \ell^\gamma) = \ell^\beta$, so $v_\ell((d^2q - 1)^2) \geq 2\beta \geq \nu + 1$. Therefore, $\Delta \equiv 0 \pmod{\ell^{\nu+1}}$, a contradiction with the definition of ν .
(2) If $\ell = 2$ and $t \equiv dq + d^{-1} \pmod{2^\nu}$ but not $\pmod{2^{\nu+1}}$ then we can write $t \equiv dq + d^{-1} + 2^\nu \pmod{2^{\nu+1}}$. We get $\Delta \equiv (dq - d^{-1})^2 \equiv 0 \pmod{2^{\nu+1}}$ nonetheless, and get a contradiction with the definition of ν .
(3) The residue class $dq + d^{-1} \pmod{\ell^{2\beta+1}}$ is well-defined by Lemma 5 because $\beta < \gamma$. If $t \equiv dq + d^{-1} \pmod{\ell^{2\beta+1}}$ then $\Delta \equiv (dq - d^{-1})^2 \pmod{\ell^{2\beta+1}}$ and thus $\nu = 2\beta$ since $(d^2q - 1, \ell^\gamma) = \ell^\beta$. \square

The third technical lemma, Lemma 8, concerns the situation where β is as large as possible, i.e. $\beta = \gamma$, and says that any solution modulo ℓ^i is a lift of the distinguished d^{-1} modulo ℓ^γ .

Lemma 8. Let $\gamma \geq 1$, $d \in (\mathbb{Z}/\ell^\gamma\mathbb{Z})^\times$, $\ell \nmid q$, and suppose that $d^2q \equiv 1 \pmod{\ell^\gamma}$ and $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma}}$, which is well-defined by Lemma 5. For any $i \geq 2\gamma$, if c is a solution to $c^2 - tc + q \equiv 0 \pmod{\ell^i}$ then $c \equiv d^{-1} \pmod{\ell^\gamma}$.

Proof. Suppose first that $\ell \neq 2$. If c_0 satisfies $c_0^2 - tc_0 + q \equiv 0 \pmod{\ell^i}$ then

$$c_0^2 - (dq + d^{-1})c_0 + q \equiv 0 \pmod{\ell^{2\gamma}}$$

because $i \geq 2\gamma$. By completing the square we find

$$(2c_0 - (dq + d^{-1}))^2 \equiv (dq - d^{-1})^2 \equiv 0 \pmod{\ell^{2\gamma}}.$$

We conclude $2c_0 \equiv dq + d^{-1} \pmod{\ell^\gamma}$. Since $dq \equiv d^{-1} \pmod{\ell^\gamma}$, we have $c_0 \equiv d^{-1} \pmod{\ell^\gamma}$.

Now consider the case $\ell = 2$. Since $2 \nmid q$ we have that d, d^{-1} , and q are all odd. Hence $t \equiv dq + d^{-1} \pmod{2^{2\gamma}}$ is even. Then $2^{-1}(dq + d^{-1})$ is defined modulo $2^{2\gamma-1}$, and $2^{-2}(dq + d^{-1})^2$ is defined modulo $2^{2\gamma}$. By completing the square we have

$$(c_0 - 2^{-1}(dq + d^{-1}))^2 \equiv 2^{-2}(dq - d^{-1})^2 \equiv 0 \pmod{2^{2\gamma}}.$$

Thus $c_0 \equiv 2^{-1}(dq + d^{-1}) \equiv d^{-1} \pmod{2^\gamma}$. \square

With Lemmas 6, 7, and 8 we may now evaluate $W(d)$. Recall the notation \sum^{*m} is a sum over elements $c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ that satisfy $c^2 - tq + q \equiv 0 \pmod{\ell^{\alpha+\gamma}}$ and lift to solutions of this polynomial modulo $\ell^{\alpha+\gamma+\min(\kappa, \alpha+\gamma)}$. In particular, it depends on the value of m . Also recall the definition $D(t; n) = \delta_n(t, dq + d^{-1})$ from (30). Let $S(a, n)$ be the number of solutions to $x^2 - a \equiv 0 \pmod{n}$. Later, in Lemma 10, we give an explicit evaluation for $S(a, n)$.

Lemma 9. (1) *If $\beta < \gamma$ then*

$$\sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}) = \begin{cases} D(t; \ell^{\gamma+\alpha+\min(\kappa, \alpha+\gamma)})\ell^\alpha & \text{if } \min(\kappa, \alpha + \gamma) \leq \beta - \alpha \\ D(t; \ell^{\beta+\gamma})\ell^{\beta-\min(\kappa, \alpha+\gamma)} & \text{if } \beta - \alpha \leq \min(\kappa, \alpha + \gamma). \end{cases}$$

(2) *If $\beta = \gamma$ then*

$$\sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}) = \begin{cases} D(t; \ell^{\gamma+\alpha+\min(\kappa, \alpha+\gamma)})\ell^\alpha & \text{if } \min(\kappa, \alpha + \gamma) \leq \gamma - \alpha \\ D(t; \ell^{2\gamma}) \frac{S(\Delta^*, \ell^{\alpha+\gamma+\min(\kappa, \alpha+\gamma)})}{\ell^{\min(\kappa, \alpha+\gamma)}} & \text{if } \gamma - \alpha \leq \min(\kappa, \alpha + \gamma), \end{cases}$$

where

$$\Delta^* = \begin{cases} t^2 - 4q & \ell \neq 2 \\ \frac{t^2 - 4q}{4} & \ell = 2. \end{cases}$$

Note that in the case $\min(\kappa, \alpha + \gamma) = \beta - \alpha$ both “if” statements are true, and the evaluations for $W(d)$ coincide. Note also that if $\ell = 2$ then $W(d)$ is only supported on $t \in 2\mathbb{Z}$ so that $\Delta^* \in \mathbb{Z}$, or else $W(d)$ is 0.

Proof. When $\min(\kappa, \alpha + \gamma) \leq \beta - \alpha$, the conclusion of Lemma 9 is immediate from Lemma 6 taking $B = \beta$, $C = \gamma$, $D = d$ and $i = \alpha + \gamma + \min(\kappa, \alpha + \gamma)$. For the rest of the proof we suppose that $\beta - \alpha \leq \min(\kappa, \alpha + \gamma)$.

(1) Assume that $\beta < \gamma$. Let $k = \alpha + \gamma + \min(\kappa, \alpha + \gamma)$. We have already treated the cases $\alpha + \gamma \leq k \leq \beta + \gamma$, so we may assume that $\beta + \gamma < k$. By Lemma 6 with $B = \beta$, $C = \gamma$, $D = d$ and $i = \alpha + \gamma$ we have

$$W(d) = \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}) = D(t; \ell^{\alpha+\gamma}) \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}).$$

By Lemma 7 (iii) we have that $k \leq \alpha + \beta + \gamma$ when the congruence condition above is satisfied, so we may apply Lemma 6 with $B = \beta$, $C = \alpha + \gamma$, $D = c^{-1}$ and $i = k$. Thus

$$W(d) = \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times} \delta_{\ell^k}(t, c^{-1}q + c) \delta_{\ell^\gamma}(c, d^{-1}).$$

We can simplify this further. Choose any $c_0 \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ such that $c_0 \equiv d^{-1} \pmod{\ell^\gamma}$. We parameterize the $c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ such that $c \equiv d^{-1} \pmod{\ell^\gamma}$ by $c = c_0 + i\ell^\gamma$, where $0 \leq i < \ell^\alpha$. Lemma 5 implies that the residue class $c^{-1}q + c \pmod{\ell^k}$ is well-defined. Since

$$(c_0 + i\ell^\gamma)^{-1} \equiv c_0^{-1} - c_0^{-2}i\ell^\gamma \pmod{\ell^k},$$

we have

$$(59) \quad \{c^{-1}q + c \pmod{\ell^k} : c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times \text{ and } c \equiv d^{-1} \pmod{\ell^\gamma}\} \\ = \{c_0^{-1}q + c_0 - (c_0^{-2}q - 1)i\ell^\gamma \pmod{\ell^k} : 0 \leq i < \ell^\alpha\}.$$

We have that $c_0 \equiv d^{-1} \pmod{\ell^\gamma}$, $\beta < \gamma$, $k > \beta + \gamma$, and $d^2q - 1 \equiv 0 \pmod{\ell^\beta}$, but $d^2q - 1 \not\equiv 0 \pmod{\ell^{\beta+1}}$. Therefore, there exists z with $0 \leq z < \ell^{k-(\gamma+\beta)}$ such that

$$-(c_0^{-2}q - 1)i\ell^\gamma \equiv zi\ell^{\beta+\gamma} \not\equiv 0 \pmod{\ell^k}.$$

This implies

$$W(d) = \sum_{i \in \mathbb{Z}/\ell^\alpha\mathbb{Z}} \delta_{\ell^k}(t, c_0^{-1}q + c_0 + iz\ell^{\beta+\gamma}).$$

As we vary over all i satisfying $0 \leq i < \ell^\alpha$, we see that $c_0^{-1}q + c_0 + zi\ell^{\beta+\gamma} \pmod{\ell^k}$ represents each residue classes modulo ℓ^k that is a lift of $c_0^{-1}q + c_0 \equiv dq + d^{-1} \pmod{\ell^{\beta+\gamma}}$ exactly $\ell^{\alpha-(k-\gamma-\beta)}$ times. Thus

$$W(d) = \ell^{\alpha-(k-\beta-\gamma)} \sum_{i \in \mathbb{Z}/\ell^{k-\beta-\gamma}\mathbb{Z}} \delta_{\ell^k}(t, c_0^{-1}q + c_0 + iz\ell^{\beta+\gamma}) \\ = \ell^{\beta-\min(\kappa, \alpha+\gamma)} D(t; \ell^{\beta+\gamma}).$$

- (2) Assume that $\beta = \gamma$. Let $k = \alpha + \gamma + \min(\kappa, \alpha + \gamma)$, so $\alpha + \gamma \leq k \leq 2\alpha + 2\gamma$. In fact we may assume that $2\gamma \leq k \leq 2\alpha + 2\gamma$ as the other cases have already been treated. By Lemma 6 with $B = \gamma$, $C = \gamma$, and $D = d$ we have

$$W(d) = \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}) = D(t; \ell^{2\gamma}) \sum_{c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times}^{*m} \delta_{\ell^\gamma}(c, d^{-1}).$$

We claim that for each k satisfying $2\gamma \leq k \leq 2\alpha + 2\gamma$ and for each $c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ such that $c \equiv d^{-1} \pmod{\ell^\gamma}$ there are either $\ell^{k-(\alpha+\gamma)}$ lifts \tilde{c} of c satisfying $\tilde{c}^2 - t\tilde{c} + q \equiv 0 \pmod{\ell^k}$ or 0 such lifts.

To see this, for each $c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ such that $c \equiv d^{-1} \pmod{\ell^\gamma}$ we apply Lemma 6 with $B = \gamma$, $C = \alpha + \gamma$ and $D = c^{-1} \pmod{\ell^{\alpha+\gamma}}$. For each lift $\tilde{c} \pmod{\ell^{\alpha+2\gamma}}$ of c thus produced, we apply Lemma 6 again with $B = \gamma$, $C = \alpha + 2\gamma$, and $D = \tilde{c}^{-1} \pmod{\ell^{\alpha+2\gamma}}$. We get either exactly $\ell^{k-(\alpha+\gamma)}$ lifts of c , or none, for each k satisfying $2\gamma \leq k \leq 2\alpha + 2\gamma \leq \alpha + 3\gamma$.

Completing the square shows that there are $S(\Delta^*, \ell^k)$ total solutions to $c^2 - tc + q \equiv 0 \pmod{\ell^k}$. Note that if $\ell = 2$ we may assume t is even because of the factor $\delta_{\ell^{2\gamma}}(t, dq + d^{-1})$, so we have that $\Delta^* \in \mathbb{Z}$, and completing the square makes sense. By Lemma 8, all solutions to $c^2 - tc + q \equiv 0 \pmod{\ell^k}$ with $k \geq 2\gamma$ have $c \equiv d^{-1} \pmod{\ell^\gamma}$. By the previous claim, this implies that exactly

$$\frac{S(\Delta^*, \ell^k)}{\ell^{k-(\alpha+\gamma)}}$$

of the $c \in (\mathbb{Z}/\ell^{\alpha+\gamma}\mathbb{Z})^\times$ with $c \equiv d^{-1} \pmod{\ell^\gamma}$ have lifts to solutions modulo ℓ^k . □

The next lemma gives an evaluation of $S(a, n)$. It is a special case of more general results in several variables going back at least to Jordan [17].

Lemma 10. *For p an odd prime we have*

$$S(a, p) = 1 + \left(\frac{a}{p}\right).$$

For an odd prime power p^ϵ with $\epsilon \geq 2$ we have

$$S(a, p^\epsilon) = \begin{cases} S(a, p) & \text{if } v_p(a) = 0 \\ 0 & \text{if } v_p(a) = 1 \\ pS(a/p^2, p^{\epsilon-2}) & \text{if } v_p(a) \geq 2. \end{cases}$$

If $p = 2$ we have

$$S(a, 2) = 1$$

$$S(a, 4) = \begin{cases} 2 & \text{if } a \equiv 0, 1 \pmod{4} \\ 0 & \text{if } a \equiv 2, 3 \pmod{4} \end{cases}$$

and if $\epsilon \geq 3$ then

$$S(a, 2^\epsilon) = \begin{cases} 4\delta_8(a, 1) & \text{if } v_2(a) = 0 \\ 0 & \text{if } v_2(a) = 1 \\ 2S(a/4, 2^{\epsilon-2}) & \text{if } v_2(a) \geq 2. \end{cases}$$

Proof Sketch. Clear factors of p common to a and p^ϵ , and apply Hensel's lemma. \square

The next lemma is a standard result relating class numbers of orders to the class numbers of the maximal orders containing them. For a proof, see for example [7, Corollary 7.28].

Lemma 11. *For $d < 0$, $d \equiv 0, 1 \pmod{4}$ and $f \in \mathbb{N}$ we have*

$$(60) \quad h_w(f^2d) = h_w(d) f \prod_{p|f} \left(1 - \left(\frac{d}{p}\right) \frac{1}{p}\right).$$

In particular, if f is a prime power dividing d then

$$(61) \quad h_w(f^2d) = h_w(d) f$$

and if $(d, f) = 1$ and d is a square modulo f (resp. $4f$ if $2 | f$) then

$$(62) \quad h_w(f^2d) = h_w(d)\varphi(f).$$

The last lemma gives identities between class numbers that we will use later.

Lemma 12. *Let $0 \leq \alpha \leq \beta \leq \gamma$ and $\ell \nmid q$ prime. Assume $(d^2q - 1, \ell^\gamma) = \ell^\beta$. If $\beta < \gamma$, then*

$$\begin{aligned} & 2 \sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} H_{\ell^\gamma, \ell^{\alpha+j}}(t, q, d) \\ &= H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma}) + \sum_{j=1}^{\beta-\alpha} H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) (\ell^j D(t; \ell^{\alpha+\gamma+j}) - \ell^{j-1} D(t; \ell^{\alpha+\gamma+j-1})) \end{aligned}$$

and if $\beta = \gamma$ we have

$$2 \sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} H_{\ell^\gamma, \ell^{\alpha+j}}(t, q, d) = H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma}) \\ + \sum_{j=1}^{\gamma-\alpha-1} H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) (\ell^j D(t; \ell^{\alpha+\gamma+j}) - \ell^{j-1} D(t; \ell^{\alpha+\gamma+j-1})) + \ell^{\gamma-\alpha} D(t; \ell^{2\gamma}) H\left(\frac{\Delta}{\ell^{2\gamma}}\right).$$

Proof. The expression in the lemma is by definition

$$\sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} \left[H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) \delta_{\ell^{\alpha+j}}(1, d^2 q) D(t; \ell^{\alpha+\gamma+j}) \right. \\ \left. - \sum_{k=1}^{\gamma-(\alpha+j)-1} H\left(\frac{\Delta}{\ell^{2(\alpha+j+k)}}\right) \delta_{\ell^{\alpha+k+j}}(1, d^2 q) (D(t; \ell^{\alpha+\gamma+j-k-1}) - D(t; \ell^{\alpha+\gamma+j+k})) \right].$$

Note that $\delta_{\ell^{\alpha+j}}(1, d^2 q) = 1$ since this is equivalent to $j \leq \beta - \alpha$. The term $\delta_{\ell^{\alpha+k+j}}(1, d^2 q)$ is equal to 1 if and only if $j + k \leq \beta - \alpha$. Swapping order of summation and writing $j + k = i$ this is

$$\sum_{j=0}^{\beta-\alpha} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) D(t; \ell^{\alpha+\gamma+j}) \\ - \sum_{i=1}^{\min(\gamma-\alpha-1, \beta-\alpha)} \left(\sum_{j=0}^{i-1} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} \right) H\left(\frac{\Delta}{\ell^{2(\alpha+i)}}\right) (D(t; \ell^{\alpha+\gamma+i-1}) - D(t; \ell^{\alpha+\gamma+i})).$$

It is straightforward to show that

$$\sum_{j=0}^{i-1} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma-\alpha-j})}{\varphi(\ell^{\gamma-\alpha})} = \begin{cases} \ell^{i-1} & \text{if } i-1 < \gamma-\alpha \\ \psi(\ell^{\gamma-\alpha}) & \text{if } i-1 = \gamma-\alpha. \end{cases}$$

We now consider two cases. If $\beta < \gamma$ we have $\beta - \alpha \leq \gamma - \alpha - 1$, so combining terms gives

$$H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma}) + \sum_{j=1}^{\beta-\alpha} H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) (\ell^j D(t; \ell^{\alpha+\gamma+j}) - \ell^{j-1} D(t; \ell^{\alpha+\gamma+j-1})).$$

On the other hand if $\beta = \gamma$ the expression is

$$H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma}) + \sum_{j=1}^{\gamma-\alpha-1} H\left(\frac{\Delta}{\ell^{2(\alpha+j)}}\right) (\ell^j D(t; \ell^{\alpha+\gamma+j}) - \ell^{j-1} D(t; \ell^{\alpha+\gamma+j-1})) \\ + \ell^{\gamma-\alpha} D(t; \ell^{2\gamma}) H\left(\frac{\Delta}{\ell^{2\gamma}}\right).$$

□

4.5.2. *Evaluating $C(t, q, d)$.* The proof of Proposition 4 breaks into three main cases: ($\beta < \gamma$), ($\beta = \gamma$ and $\nu < 2(\alpha + \gamma)$), and ($\beta = \gamma$ and $\nu \geq 2(\alpha + \gamma)$).

The case $\beta < \gamma$.

We further split into three cases according to the value of κ . The ranges $\kappa < \beta - \alpha$, $\beta - \alpha \leq \kappa \leq \beta$, and $\kappa > \beta$ will each be treated differently.

The case $\beta < \gamma$ and $\kappa > \beta$.

We have $\beta - \alpha \leq \beta < \kappa$ and $\beta - \alpha \leq \alpha + \gamma$ so that the second case of Lemma 9 (i) applies. For such κ we have

$$(63) \quad c_\kappa(t, q, d) = \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma-\min(\kappa, \alpha+\gamma)})} \ell^{\beta-\min(\kappa, \alpha+\gamma)} D(t; \ell^{\beta+\gamma}).$$

Since $2\beta < 2\kappa \leq \nu$ and $2\beta + 1 \leq \beta + \gamma$, Lemma 7 (iii) applies, and so $D(t; \ell^{\beta+\gamma}) = 0$. We conclude that $c_\kappa(t, q, d) = 0$ for all $\kappa > \beta$.

The case $\beta < \gamma$ and $\beta - \alpha \leq \kappa \leq \beta$.

We have $\min(\kappa, \alpha + \gamma) = \kappa$ since $\kappa \leq \beta < \gamma \leq \alpha + \gamma$. The second case of Lemma 9 (i) applies and we again have

$$c_\kappa(t, q, d) = \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma-\kappa})} \ell^{\beta-\kappa} D(t; \ell^{\beta+\gamma}).$$

Suppose that $\nu \neq 2\beta$. Because $2\beta + 1 \leq \beta + \gamma$, Lemma 7 (iii) again applies and we see that $D(t; \ell^{\beta+\gamma}) = 0$. Therefore for any $\beta - \alpha \leq \kappa \leq \beta$ with $\beta < \gamma$ we have

$$(64) \quad c_\kappa(t, q, d) = \delta(\nu, 2\beta) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\psi(\ell^{\alpha+\gamma-\kappa})} \ell^{\beta-\kappa} D(t; \ell^{\beta+\gamma}).$$

The case $\beta < \gamma$ and $\kappa < \beta - \alpha$.

In this range of κ we again have $\min(\kappa, \alpha + \gamma) = \kappa < \beta - \alpha$ since $\kappa < \beta - \alpha \leq \alpha + \gamma$. The first case of Lemma 9 (i) then applies and we have for $\kappa < \beta - \alpha$ that

$$(65) \quad c_\kappa(t, q, d) = \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \ell^{\alpha+\kappa} D(t; \ell^{\alpha+\gamma+\kappa}).$$

We would like to apply Lemma 11 to the class numbers occurring in (65), but this requires a certain hypothesis to hold. We verify this hypothesis with the following lemma.

Lemma 13. *Suppose that $t \equiv dq + d^{-1} \pmod{\ell^{\alpha+\gamma+\kappa}}$, $\beta < \gamma$ and $\kappa < \beta - \alpha$. Then*

$$v_\ell \left(\frac{\Delta}{\ell^{2\alpha} m^2} \right) \geq 2,$$

and

$$\frac{\Delta}{\ell^{2\alpha+2} m^2} \equiv 0, 1 \pmod{4}.$$

Proof. Since $\kappa < \beta - \alpha$, if $\nu = 2\beta$ then $v_\ell(\Delta/(\ell^{2\alpha} m^2)) = 2\beta - 2\alpha - 2\kappa \geq 2$. So suppose that $\nu \neq 2\beta$. Then by Lemma 7 (i) and (iii) we have that $\alpha + \gamma + \kappa \leq \min(\nu, 2\beta)$. Thus

$$v_\ell \left(\frac{\Delta}{\ell^{2\alpha} m^2} \right) = \nu - 2\alpha - 2\kappa \geq 2\gamma + \nu - 2 \min(\nu, 2\beta).$$

If $2\beta < \nu$ then

$$2\gamma + \nu - 2 \min(\nu, 2\beta) = \nu + 2\gamma - 4\beta > 2(\gamma - \beta) \geq 2.$$

If $\nu < 2\beta$ then

$$2\gamma + \nu - 2 \min(\nu, 2\beta) = 2\gamma - \nu > 2(\gamma - \beta) \geq 2.$$

If $\ell \neq 2$ then $\ell^2 \equiv 1 \pmod{4}$ so the second statement of the lemma follows immediately from the first. Now suppose $\ell = 2$. There are two cases, $0 \leq \kappa \leq \beta - \alpha - 2$, and $\kappa = \beta - \alpha - 1$.

In the case $\kappa \leq \beta - \alpha - 2$, if $\nu \geq 2\beta$ then

$$\nu - 2\alpha - 2\kappa \geq 2\beta - 2\alpha - 2\kappa \geq 4$$

so that $\Delta/(\ell^{2\alpha} m^2) \equiv 0 \pmod{4}$. If $\nu < 2\beta$ then by Lemma 7 (ii) we have $\alpha + \gamma + \kappa \leq \nu - 1$, and so $\nu - 2\alpha - 2\kappa \geq 4$ so that $\Delta/(\ell^{2\alpha} m^2) \equiv 0 \pmod{4}$ as well.

Now suppose that $\kappa = \beta - \alpha - 1$. In this case, by assumption $t \equiv dq + d^{-1} \pmod{2^{\beta+\gamma-1}}$. We calculate that

$$\Delta \equiv (dq - d^{-1})^2 \pmod{2^{\beta+\gamma+1}}$$

so by definition of β and the fact that $\beta < \gamma$ we have that $\Delta/2^{2\beta} = \Delta/\ell^{2(\alpha+\kappa+1)}$ is a square modulo 4, as was to be shown. \square

By the first part of Lemma 13, we may apply Lemma 11 to (65) with $f = \ell^\alpha$ and $d = \Delta/(\ell^{2\alpha} m^2)$, which is divisible by ℓ . Thus for $\beta < \gamma$ and $\kappa < \beta - \alpha$ we have

$$c_\kappa(t, q, d) = \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{2\alpha} m^2} \right) \ell^{2\alpha+\kappa} D(t; \ell^{\alpha+\gamma+\kappa}).$$

We now assemble the three cases for the ranges of κ . The definition of the Hurwitz-Kronecker class number H from (28) implies that if $\nu \geq 2(\kappa + \alpha)$, then

$$(66) \quad \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{2\alpha} m^2} \right) = \begin{cases} H \left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}} \right) - H \left(\frac{\Delta}{\ell^{2(\alpha+\kappa+1)}} \right) & \text{if } \nu \geq 2(\alpha + \kappa + 1) \text{ and } \frac{\Delta}{\ell^{2(\alpha+\kappa+1)}} \equiv 0, 1 \pmod{4} \\ H \left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}} \right) & \text{otherwise.} \end{cases}$$

By the second part of Lemma 13 the conditions on the first line of (66) hold. So when $\beta < \gamma$ we have that

$$\begin{aligned}
C(t, q, d) &= \sum_{\kappa \geq 0} c_\kappa(t, q, d) \\
&= \sum_{\kappa=0}^{\beta-\alpha-1} \left(H\left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}}\right) - H\left(\frac{\Delta}{\ell^{2(\alpha+\kappa+1)}}\right) \right) D(t; \ell^{\alpha+\gamma+\kappa}) \ell^{2\alpha+\kappa} \\
&\quad + \delta(\nu, 2\beta) D(t; \ell^{\beta+\gamma}) \ell^\beta \sum_{\kappa=\beta-\alpha}^{\beta} \sum_{\substack{m^2 | \Delta \\ v_\ell(m)=\kappa}} h_w\left(\frac{\Delta}{m^2}\right).
\end{aligned}$$

Rearranging terms according to the argument of H we have

$$\begin{aligned}
(67) \quad C(t, q, d) &= H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma}) \ell^{2\alpha} \\
&\quad - \sum_{\kappa=1}^{\beta-\alpha-1} H\left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}}\right) (\ell^{2\alpha+\kappa-1} D(t; \ell^{\alpha+\gamma+\kappa-1}) - \ell^{2\alpha+\kappa} D(t; \ell^{\alpha+\gamma+\kappa})) \\
&\quad - H\left(\frac{\Delta}{\ell^{2\beta}}\right) D(t; \ell^{\beta+\gamma-1}) \ell^{\alpha+\beta-1} \\
&\quad + \delta(\nu, 2\beta) D(t; \ell^{\beta+\gamma}) \ell^\beta \sum_{\kappa=\beta-\alpha}^{\beta} \sum_{\substack{m^2 | \Delta \\ v_\ell(m)=\kappa}} h_w\left(\frac{\Delta}{m^2}\right).
\end{aligned}$$

We claim that the last line of (67) is equal to

$$\delta(\nu, 2\beta) D(t; \ell^{\beta+\gamma}) \ell^{\beta+\alpha} H\left(\frac{\Delta}{\ell^{2\beta}}\right).$$

Suppose that $\Delta/\ell^{2\beta} \in \mathbb{Z}$ and $t \equiv dq + d^{-1} \pmod{\ell^{\beta+\gamma}}$, since otherwise the last line of (67) vanishes. Then

$$\Delta \equiv (dq - d^{-1})^2 \pmod{\ell^{\beta+\gamma}},$$

and since $\beta < \gamma$ we have that $\Delta/\ell^{2\beta}$ is a square modulo ℓ . If $\ell = 2$ then since $\ell \nmid q$ we have

$$\Delta \equiv (dq - d^{-1})^2 \pmod{2^{\beta+\gamma+2}}$$

so that $\Delta/2^{2\beta}$ is a square modulo 8. Then Lemma 11 applies with $f = \ell^{\beta-\kappa}$, and $d = \Delta/\ell^{2\beta}$, which is a square modulo ℓ (resp. 8). So in the case $\beta - \alpha \leq \kappa \leq \beta$ we have

$$(68) \quad \sum_{\substack{m^2 | \Delta \\ v_\ell(m)=\kappa}} h_w\left(\frac{\Delta}{m^2}\right) = \sum_{\substack{m^2 | \Delta \\ v_\ell(m)=\kappa}} h_w\left(\frac{\Delta}{\ell^{2(\beta-\kappa)} m^2}\right) \varphi(\ell^{\beta-\kappa}).$$

Therefore the last line of (67) simplifies as

$$\begin{aligned}
(69) \quad & \delta(\nu, 2\beta)D(t; \ell^{\beta+\gamma})\ell^\beta \sum_{\kappa=\beta-\alpha}^{\beta} \sum_{\substack{m^2|\Delta \\ v_\ell(m)=\kappa}} h_w \left(\frac{\Delta}{m^2} \right) \\
& = \delta(\nu, 2\beta)D(t; \ell^{\beta+\gamma})\ell^\beta \sum_{\kappa=\beta-\alpha}^{\beta} \sum_{\substack{m^2|\Delta \\ v_\ell(m)=\kappa}} h_w \left(\frac{\Delta}{\ell^{2(\beta-\kappa)}m^2} \right) \varphi(\ell^{\beta-\kappa}).
\end{aligned}$$

If this term does not vanish we have $\nu = 2\beta$ and so

$$\sum_{\substack{m^2|\Delta/\ell^{2\beta} \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^{2\beta}m^2} \right) = H \left(\frac{\Delta}{\ell^{2\beta}} \right)$$

by the definition of Hurwitz-Kronecker class numbers (28). Therefore the expression in (69) equals

$$(70) \quad \delta(\nu, 2\beta)D(t; \ell^{\beta+\gamma})\ell^\beta H \left(\frac{\Delta}{\ell^{2\beta}} \right) \sum_{\kappa=\beta-\alpha}^{\beta} \varphi(\ell^{\beta-\kappa}) = \delta(\nu, 2\beta)D(t; \ell^{\beta+\gamma})\ell^{\beta+\alpha} H \left(\frac{\Delta}{\ell^{2\beta}} \right),$$

proving the claim above.

Putting this back into (67), when $\beta < \gamma$ we have

$$\begin{aligned}
(71) \quad & C(t, q, d) = \ell^{2\alpha} H \left(\frac{\Delta}{\ell^{2\alpha}} \right) \delta_{\ell^{\alpha+\gamma}}(t, dq + d^{-1}) \\
& - \ell^{2\alpha} \sum_{\kappa=1}^{\beta-\alpha} H \left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}} \right) (\ell^{\kappa-1} \delta_{\ell^{\alpha+\gamma+\kappa-1}}(t, dq + d^{-1}) - \ell^\kappa \delta_{\ell^{\alpha+\gamma+\kappa}}(t, dq + d^{-1})).
\end{aligned}$$

Finally, by applying Lemma 12 we conclude that

$$C(t, q, d) = 2\ell^{2\alpha} \sum_{j=0}^{\beta-\alpha} \varphi(\ell^j) H_{\ell^\gamma, \ell^{\alpha+j}, d}(t, q),$$

which in the case $\beta < \gamma$ matches exactly the expression in Proposition 4.

The case $\beta = \gamma$ and $\nu < 2(\alpha + \gamma)$.

Recall the definition of $c_\kappa(t, q, d)$ from (57), so that we have

$$C(t, q, d) = \sum_{\kappa \geq 0} c_\kappa(t, q, d).$$

We split into three cases according to the value of κ : each of the ranges $\kappa < \gamma - \alpha$, $\gamma - \alpha \leq \kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$, and $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$ will be treated differently. Specifically we

write

$$\text{I} \stackrel{\text{def}}{=} \sum_{\kappa < \gamma - \alpha} c_\kappa(t, q, d), \quad \text{II} \stackrel{\text{def}}{=} \sum_{\kappa = \gamma - \alpha}^{2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha) - 1} c_\kappa(t, q, d), \quad \text{III} \stackrel{\text{def}}{=} \sum_{\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)} c_\kappa(t, q, d)$$

so that $C(t, q, d) = \text{I} + \text{II} + \text{III}$. Note that the second range of κ above has been chosen to have an even number of terms.

The case $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha) \geq \gamma - \alpha$, i.e. the sum III.

We have $\kappa \leq \lfloor \nu/2 \rfloor < \alpha + \gamma$ by hypothesis, so $\min(\kappa, \alpha + \gamma) = \kappa$. Therefore the second case of Lemma 9 (ii) applies and we have for $\beta = \gamma$, $\nu < 2(\alpha + \gamma)$ and $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$ that

$$(72) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) S(\Delta^*, \ell^{\alpha + \gamma + \kappa}),$$

where we note that if $\ell = 2$ then $\Delta^* \in \mathbb{Z}$ otherwise $c_\kappa(t, q, d) = 0$. A repeated application of Lemma 10 shows that when $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$ we have that

$$(73) \quad S(\Delta^*, \ell^{\alpha + \gamma + \kappa}) = \begin{cases} \ell^{\nu/2} & \nu \text{ even, } \kappa = \nu - (\alpha + \gamma) \text{ and } \Delta/\ell^\nu \equiv 0, 1 \pmod{4} \\ \ell^{\nu/2} \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right) & \nu \text{ even, } \kappa > \nu - (\alpha + \gamma) \text{ and } \Delta/\ell^\nu \equiv 0, 1 \pmod{4} \\ 0 & \nu \text{ odd, or } \Delta/\ell^\nu \equiv 2, 3 \pmod{4}. \end{cases}$$

If $t \equiv dq + d^{-1} \pmod{\ell^{\alpha + \gamma + \kappa}}$ then Lemma 7 (i) implies that $\nu \geq 2\gamma$. Therefore for $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$ we have that:

(1) If ν even, $\kappa = \nu - (\alpha + \gamma)$, $2\gamma \leq \nu$, and $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ then

$$c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\nu/2} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right).$$

(2) If ν even, $\kappa > \nu - (\alpha + \gamma)$, $2\gamma \leq \nu$, and $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ then

$$c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\nu/2} \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right).$$

(3) If ν odd, or $2\gamma < \nu$, or $\Delta/\ell^\nu \equiv 2, 3 \pmod{4}$ then $c_\kappa(t, q, d) = 0$.

If ν is even, $\kappa = \nu - (\alpha + \gamma)$, $2\gamma \leq \nu$ and $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ then we may apply Lemma 11 with $d = \Delta/(\ell^{\nu+2\kappa}m^2)$ and $f = \ell^{\nu/2-\kappa}$ to find

$$\begin{aligned}
(74) \quad c_\kappa(t, q, d) &= D(t; \ell^{2\gamma}) \ell^{\nu-\kappa} \sum_{\substack{m^2|\Delta \\ v_\ell(m)=\kappa}} h_w \left(\frac{\Delta}{\ell^{\nu-2\kappa}m^2} \right) \left(1 - \left(\frac{\Delta}{\ell^{\nu-2\kappa}m^2} \right) \frac{1}{\ell} \right) \\
&= D(t; \ell^{2\gamma}) \ell^{\nu-\kappa} \sum_{\substack{m^2|(\Delta/\ell^\nu) \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^\nu m^2} \right) \left(1 - \left(\frac{\Delta}{\ell^\nu m^2} \right) \frac{1}{\ell} \right) \\
&= D(t; \ell^{2\gamma}) \ell^{\nu-\kappa} \left(1 - \left(\frac{\Delta/\ell^\nu}{\ell} \right) \frac{1}{\ell} \right) \sum_{\substack{m^2|(\Delta/\ell^\nu) \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^\nu m^2} \right),
\end{aligned}$$

where the second equality follows from the change of variables $m \rightarrow m\ell^\kappa$ and the third equality follows from the fact that $m \not\equiv 0 \pmod{\ell}$, so $\left(\frac{m^2}{\ell}\right) = 1$.

Similarly, in the case that ν is even, $\nu - (\alpha + \gamma) < \kappa < \nu/2$, $2\gamma \leq \nu$ and $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ we have that

$$(75) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\nu-\kappa} \left(1 - \left(\frac{\Delta/\ell^\nu}{\ell} \right) \frac{1}{\ell} \right) \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right) \sum_{\substack{m^2|(\Delta/\ell^\nu) \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^\nu m^2} \right).$$

If ν is even, $\kappa = \nu/2$, $2\gamma \leq \nu$ and $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ then we do not apply Lemma 11 to $c_\kappa(t, q, d)$, but change variables $m \rightarrow m\ell^\kappa$. We have

$$(76) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\nu/2} \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right) \sum_{\substack{m^2|(\Delta/\ell^\nu) \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^\nu m^2} \right).$$

Observe that the following is a telescoping sum:

$$\begin{aligned}
(77) \quad \psi(\ell^{\alpha+\gamma}) &= \ell^{\alpha+\gamma} \left(1 - \left(\frac{\Delta/\ell^\nu}{\ell} \right) \frac{1}{\ell} \right) + \sum_{\kappa=\nu-(\alpha+\gamma)+1}^{\nu/2-1} \ell^{\nu-\kappa} \left(1 - \left(\frac{\Delta/\ell^\nu}{\ell} \right) \frac{1}{\ell} \right) \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right) \\
&\quad + \ell^{\nu/2} \left(1 + \left(\frac{\Delta/\ell^\nu}{\ell} \right) \right).
\end{aligned}$$

Thus, taking the sum over $\kappa \geq 2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma - \alpha)$ of the expressions (74), (75), and (76), we have by (77) that if $\beta = \gamma$ and $\nu < 2(\alpha + \gamma)$, then

$$\text{III} = \begin{cases} D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{\substack{m^2|(\Delta/\ell^\nu) \\ v_\ell(m)=0}} h_w \left(\frac{\Delta}{\ell^\nu m^2} \right) & \text{if } \nu \text{ is even and } \nu \geq 2\gamma \\ 0 & \text{if } \nu \text{ is odd or } \nu < 2\gamma. \end{cases}$$

Note that the condition $\Delta/\ell^\nu \equiv 0, 1 \pmod{4}$ is now implicit in the definition of h_w .

The case $\gamma - \alpha \leq \kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$, i.e. the sum II.

We have $\kappa \leq \lfloor \nu/2 \rfloor < \alpha + \gamma$ by hypothesis, so $\min(\kappa, \alpha + \gamma) = \kappa$. Therefore the second case of Lemma 9 (ii) applies, so for $\beta = \gamma$, $\nu < 2(\alpha + \gamma)$ and $\gamma - \alpha \leq \kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$ we have

$$(78) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) S(\Delta^*, \ell^{\alpha+\gamma+\kappa}).$$

Now, for $\gamma - \alpha \leq \kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$ we have by Lemma 10 that

$$(79) \quad S(\Delta^*, \ell^{\alpha+\gamma+\kappa}) = \begin{cases} \ell^{\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} & \text{if } \Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 0, 1 \pmod{4} \\ 0 & \text{if } \Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 2, 3 \pmod{4}. \end{cases}$$

If $t \equiv dq + d^{-1} \pmod{\ell^{\alpha+\gamma+\kappa}}$ then Lemma 7 (i) implies that $\nu \geq 2\gamma$. Therefore if $\Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 0, 1 \pmod{4}$ and $\nu \geq 2\gamma$ we have for $\gamma - \alpha \leq \kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$ that

$$(80) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \ell^{\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor},$$

and if $\nu < 2\gamma$ or $\Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 2, 3 \pmod{4}$ then $c_\kappa(t, q, d) = 0$.

We simplify $c_\kappa(t, q, d)$ in two cases according to whether $\alpha + \gamma + \kappa$ is even or odd. First, assume that $\alpha + \gamma + \kappa$ is even. Then the assumption $\kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$ implies that $\nu - (\alpha + \gamma + \kappa) > 0$. If $\Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 0, 1 \pmod{4}$ we apply Lemma 11 with $d = \frac{\Delta}{\ell^{\alpha+\gamma-\kappa} m^2}$ and $f = \ell^{\alpha+\gamma - (\frac{\alpha+\gamma+\kappa}{2})}$. Therefore

$$(81) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\alpha+\gamma} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{\alpha+\gamma-\kappa} m^2} \right).$$

Now suppose $\alpha + \gamma + \kappa$ is odd. The assumption $\kappa < 2\lfloor \frac{\nu+1}{2} \rfloor - (\alpha + \gamma)$ implies that $\nu - (\alpha + \gamma + \kappa - 1) > 0$. If $\Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 0, 1 \pmod{4}$ we apply Lemma 11 with $d = \frac{\Delta}{\ell^{\alpha+\gamma-\kappa-1} m^2}$ and $f = \ell^{\alpha+\gamma - (\frac{\alpha+\gamma+\kappa+1}{2})}$. Thus

$$(82) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \ell^{\alpha+\gamma-1} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{\alpha+\gamma-\kappa-1} m^2} \right).$$

Putting the even and odd cases for $\alpha + \gamma + \kappa$ together, we see that if $\alpha + \gamma + \kappa \equiv 0 \pmod{2}$ and $2\gamma \leq \nu$, then

$$(83) \quad c_\kappa(t, q, d) + c_{\kappa+1}(t, q, d) = D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{\alpha+\gamma-\kappa} m^2} \right),$$

and if $\nu < 2\gamma$ then $c_\kappa(t, q, d) = 0$.

Pairing up the terms in the sum II two at a time, we have that if $2\gamma \leq \nu$

$$\begin{aligned}
(84) \quad \text{II} &= \sum_{\kappa=\gamma-\alpha}^{2\lfloor \frac{\nu+1}{2} \rfloor - (\gamma-\alpha) - 1} c_\kappa(t, q, d) \\
&= \sum_{j=0}^{\lfloor \frac{\nu+1}{2} \rfloor - \gamma - 1} (c_{\gamma-\alpha+2j}(t, q, d) + c_{\gamma-\alpha+2j+1}(t, q, d)) \\
&= D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{j=0}^{\lfloor \frac{\nu+1}{2} \rfloor - \gamma - 1} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \gamma - \alpha + 2j}} h_w \left(\frac{\Delta}{\ell^{\alpha+\gamma - (\gamma - \alpha + 2j)} m^2} \right) \\
&= D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{j=0}^{\lfloor \frac{\nu+1}{2} \rfloor - \gamma - 1} \sum_{\substack{m^2 | (\Delta / \ell^{2\gamma}) \\ v_\ell(m) = j}} h_w \left(\frac{\Delta}{\ell^{2\gamma} m^2} \right),
\end{aligned}$$

and if $2\gamma > \nu$ then $\text{II} = 0$. The definition of H from (28) implies that if $\nu < 2(\alpha + \gamma)$, then

$$(85) \quad \text{II} + \text{III} = \begin{cases} D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) H \left(\frac{\Delta}{\ell^{2\gamma}} \right) & \text{if } 2\gamma \leq \nu \\ 0 & \text{if } \nu < 2\gamma. \end{cases}$$

The case $\kappa < \gamma - \alpha$, i.e. the sum I.

These terms are treated similarly to the ($\beta < \gamma$ and $\kappa < \beta - \alpha$) case from before but there are several details that change, so we repeat the argument.

We have that $\kappa = \min(\kappa, \alpha + \gamma)$ because $\kappa < \gamma - \alpha$. The first case of Lemma 9 (ii) holds, which implies

$$(86) \quad c_\kappa(t, q, d) = D(t; \ell^{\alpha+\gamma+\kappa}) \ell^{\alpha+\kappa} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right).$$

We want to apply Lemma 11 to this expression, but need to check a certain hypothesis.

Lemma 14. (1) If $t \equiv dq + d^{-1} \pmod{\ell^{\alpha+\gamma+\kappa}}$ and $\kappa < \gamma - \alpha$ then $v_\ell(\Delta/(\ell^{2\alpha}m^2)) \geq 1$ and $\Delta/(\ell^{2\alpha}m^2) \equiv 0, 1 \pmod{4}$.

(2) If $\kappa < \gamma - \alpha - 1$ then $v_\ell(\Delta/(\ell^{2\alpha}m^2)) \geq 2$ and $\Delta/(\ell^{2\alpha+2}m^2) \equiv 0, 1 \pmod{4}$.

Proof. (1) Suppose $\nu \geq 2\gamma$. Then we have directly that

$$\nu - 2\alpha - 2\kappa \geq \nu - 2\gamma + 2 \geq 2.$$

Now suppose $\nu < 2\gamma$. By Lemma 7 (i) we have $\alpha + \gamma + \kappa \leq \nu$, from which it follows that

$$\nu - 2\alpha - 2\kappa \geq \gamma - \alpha - \kappa \geq 1.$$

If $\ell \neq 2$ then $\ell^2 \equiv 1 \pmod{4}$, and so $\Delta/(\ell^{2\alpha}m^2) \equiv 0, 1 \pmod{4}$.

Suppose $\ell = 2$. If $\nu \geq 2\gamma$ we already have $\nu - 2\gamma - 2\kappa \geq 2$ so $\Delta/(2^{2\alpha}m^2) \equiv 0 \pmod{4}$. If $\nu < 2\gamma$ then by Lemma 7 (ii) and $\kappa < \gamma - \alpha$ we have that $\nu - 2\alpha - 2\kappa \geq 2$, and so $\Delta/(2^{2\alpha}m^2) \equiv 0 \pmod{4}$ as well.

(2) We now consider the stronger condition that $\kappa \leq \gamma - \alpha - 2$. This set of κ is non-empty only if $\gamma \geq 2$ so we assume this. If $\nu \geq 2\gamma$ we have directly

$$\nu - 2\alpha - 2\kappa \geq \nu - 2\gamma + 2 \geq 2$$

as above. If $\nu < 2\gamma$ then we have

$$\nu - 2\alpha - 2\kappa \geq \gamma - \alpha - \kappa \geq 2$$

by Lemma 7 (i). If $\ell \neq 2$ then $\ell^2 \equiv 1 \pmod{4}$ so we have $\Delta/(\ell^{2(\alpha+1)}m^2) \equiv 0, 1 \pmod{4}$.

If $\ell = 2$ then we write $t \equiv dq + d^{-1} + \epsilon 2^{\alpha+\gamma+\kappa} \pmod{2^{\alpha+\gamma+\kappa+2}}$ with $\epsilon \in \mathbb{Z}/4\mathbb{Z}$. Then because $\gamma \geq 2$ and q is odd we have that

$$\Delta \equiv (dq - d^{-1})^2 \pmod{2^{\alpha+\gamma+\kappa+2}}.$$

Since

$$\alpha + \gamma + \kappa + 2 - 2(\alpha + \kappa + 1) \geq \gamma - \alpha - \kappa \geq 2$$

we have that $\Delta/2^{2(\alpha+\kappa+1)}$ is a square modulo 4, as was to be shown. \square

By Lemma 14 we may apply Lemma 11 to $c_\kappa(t, q, d)$ with $d = \frac{\Delta}{\ell^{2\alpha}m^2}$ and $f = \ell^\alpha$. For $k < \gamma - \alpha$ this implies

$$(87) \quad c_\kappa(t, q, d) = D(t; \ell^{\alpha+\gamma+\kappa}) \ell^{2\alpha+\kappa} \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{\ell^{2\alpha}m^2} \right).$$

We apply the expression (66) to (87). If $\kappa < \gamma - \alpha - 1$ then the second part of Lemma 14 implies that

$$(88) \quad c_\kappa(t, q, d) = D(t; \ell^{\alpha+\gamma+\kappa}) \ell^{2\alpha+\kappa} \left(H \left(\frac{\Delta}{\ell^{2\alpha+2\kappa}} \right) - H \left(\frac{\Delta}{\ell^{2\alpha+2\kappa+2}} \right) \right).$$

It remains for us to consider the case $\kappa = \gamma - \alpha - 1$.

Lemma 15. *Suppose that $\beta = \gamma$. We have that $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma}}$ if and only if $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma-1}}$, $\nu \geq 2\gamma$, and $\Delta/\ell^{2\gamma} \equiv 0, 1 \pmod{4}$.*

Proof. We first prove the ‘‘only if’’ direction. Suppose that $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma}}$. Clearly $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma-1}}$, so Lemma 7 (i) directly shows that $\nu \geq 2\gamma$. If $\ell \neq 2$ then $\Delta/\ell^{2(\alpha+\kappa+1)} \equiv 0, 1 \pmod{4}$ is automatically satisfied. If $\ell = 2$ one may calculate from $t \equiv dq + d^{-1} \pmod{2^{2\gamma}}$ and the fact that q is odd that

$$\Delta \equiv (dq - d^{-1})^2 \pmod{2^{2\gamma+2}},$$

which by the assumption $\beta = \gamma$ implies that $\Delta/2^{2\gamma}$ is a square modulo 4.

We now prove the ‘‘if’’ direction. We have that $t \equiv dq + d^{-1} \pmod{\ell^{2\gamma-1}}$, which we write as $t \equiv dq + d^{-1} + \epsilon \ell^{2\gamma-1} \pmod{\ell^{2\gamma}}$ with $\epsilon \in \mathbb{Z}/\ell\mathbb{Z}$. Our goal is to show that $\epsilon = 0$.

We have that

$$\Delta \equiv (dq - d^{-1})^2 + 2\epsilon(dq + d^{-1})\ell^{2\gamma-1} \equiv 0 \pmod{\ell^{2\gamma}}.$$

Then $\beta = \gamma$ implies that $2\epsilon(dq + d^{-1}) \equiv 0 \pmod{\ell}$. We have $dq + d^{-1} \equiv 2d^{-1} \pmod{\ell}$, so if $\ell \neq 2$ we must have that $\epsilon = 0$ as was to be shown.

If $\ell = 2$ then assume $t \equiv dq + d^{-1} + \epsilon 2^{2\gamma-1} \pmod{2^{2\gamma+2}}$ where $\epsilon \in \mathbb{Z}/8\mathbb{Z}$. Our goal is to show ϵ is even. If $\gamma \geq 2$ we have by the hypothesis $\nu \geq 2\gamma$ that

$$\Delta/2^{2\gamma} \equiv \epsilon(dq + d^{-1}) \equiv 0, 1 \pmod{4}.$$

Since q is odd $dq + d^{-1} \equiv 2 \pmod{4}$, ϵ must be even and hence $t \equiv dq + d^{-1} \pmod{2^{2\gamma}}$. If $\gamma = 1$ then we have that

$$\Delta/2^{2\gamma} \equiv \epsilon(dq + d^{-1} + \epsilon) \equiv 0, 1 \pmod{4},$$

from which one checks using $dq + d^{-1} \equiv 2 \pmod{4}$ that ϵ must be even. \square

By Lemma 15 and (66), when $\kappa = \gamma - \alpha - 1$ we have that

$$(89) \quad c_\kappa(t, q, d) = D(t; \ell^{\alpha+\gamma+\kappa})\ell^{2\alpha+\kappa} H\left(\frac{\Delta}{\ell^{2\alpha+2\kappa}}\right) - D(t; \ell^{\alpha+\gamma+\kappa+1})\ell^{2\alpha+\kappa} H\left(\frac{\Delta}{\ell^{2\alpha+2\kappa+2}}\right).$$

Now putting together the formulas (88) and (89) and rearranging according to the argument of H we have that

$$\begin{aligned} \text{I} &= \sum_{\kappa < \gamma - \alpha} c_\kappa(t, q, d) \\ &= H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma})\ell^{2\alpha} \\ &\quad - \sum_{\kappa=1}^{\gamma-\alpha-1} H\left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}}\right) (\ell^{2\alpha+\kappa-1} D(t; \ell^{\alpha+\gamma+\kappa-1}) - \ell^{2\alpha+\kappa} D(t; \ell^{\alpha+\gamma+\kappa})) \\ &\quad - H\left(\frac{\Delta}{\ell^{2\gamma}}\right) D(t; \ell^{2\gamma})\ell^{\gamma+\alpha-1}. \end{aligned}$$

Adding this together with the result for II + III from (85) we find that

$$\begin{aligned} (90) \quad C(t, q, d) &= \text{I} + \text{II} + \text{III} \\ &= H\left(\frac{\Delta}{\ell^{2\alpha}}\right) D(t; \ell^{\alpha+\gamma})\ell^{2\alpha} \\ &\quad - \sum_{\kappa=1}^{\gamma-\alpha-1} H\left(\frac{\Delta}{\ell^{2(\alpha+\kappa)}}\right) (\ell^{2\alpha+\kappa-1} D(t; \ell^{\alpha+\gamma+\kappa-1}) - \ell^{2\alpha+\kappa} D(t; \ell^{\alpha+\gamma+\kappa})) \\ &\quad + H\left(\frac{\Delta}{\ell^{2\gamma}}\right) D(t; \ell^{2\gamma})\ell^{\gamma+\alpha}. \end{aligned}$$

By Lemma 12 this matches exactly the claimed formula from Proposition 4.

The case $\beta = \gamma$ and $\nu \geq 2(\alpha + \gamma)$.

Recall the definition of $c_\kappa(t, q, d)$ from (57). We split into three cases according to the value of κ : each of the ranges $\kappa < \gamma - \alpha$, $\gamma - \alpha \leq \kappa < \gamma + \alpha$, and $\kappa \geq \gamma + \alpha$ will be treated differently. Specifically we write

$$\text{I} \stackrel{\text{def}}{=} \sum_{\kappa < \gamma - \alpha} c_\kappa(t, q, d), \quad \text{II} \stackrel{\text{def}}{=} \sum_{\kappa = \gamma - \alpha}^{\gamma + \alpha - 1} c_\kappa(t, q, d), \quad \text{III} \stackrel{\text{def}}{=} \sum_{\kappa \geq \gamma + \alpha} c_\kappa(t, d)$$

so that $C(t, q, d) = \text{I} + \text{II} + \text{III}$. Note that the second range of κ above has been chosen to have an even number of terms.

The case $\kappa \geq \gamma + \alpha$, i.e. the sum III.

In this case we have $\gamma - \alpha \leq \gamma + \alpha = \min(\kappa, \alpha + \gamma)$ so the second case of Lemma 9 (ii) applies. We have that

$$(91) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) \frac{\psi(\ell^{\alpha+\gamma})}{\ell^{\alpha+\gamma}} S(\Delta^*, \ell^{2(\alpha+\gamma)}).$$

By assumption $\nu \geq 2(\alpha + \gamma)$, which implies that $\Delta^* \equiv 0 \pmod{\ell^{2(\alpha+\gamma)}}$. For $\kappa \geq \gamma + \alpha$, Lemma 10 implies that

$$S(\Delta^*, \ell^{2(\alpha+\gamma)}) = S(0, \ell^{2(\alpha+\gamma)}) = \ell^{\alpha+\gamma}.$$

Thus for $\kappa \geq \gamma + \alpha$, $\beta = \gamma$, and $\nu \geq 2(\alpha + \gamma)$ we have

$$(92) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right),$$

and therefore,

$$(93) \quad \text{III} = \sum_{\kappa \geq \alpha + \gamma} c_\kappa(t, q, d) = \sum_{j \geq \alpha} c_{\gamma+j}(t, q, d) = D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{j \geq \alpha} \sum_{\substack{m^2 | (\Delta/\ell^{2\gamma}) \\ v_\ell(m) = j}} h_w \left(\frac{\Delta}{\ell^{2\gamma} m^2} \right).$$

The case $\gamma - \alpha \leq \kappa < \gamma + \alpha$ i.e. the sum II.

In this case we have $\kappa = \min(\kappa, \alpha + \gamma)$ so the second case of Lemma 9 (ii) applies. This implies

$$(94) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2 | \Delta \\ v_\ell(m) = \kappa}} h_w \left(\frac{\Delta}{m^2} \right) S(\Delta^*, \ell^{\alpha+\gamma+\kappa}).$$

By our assumption that $\nu \geq 2(\alpha + \gamma)$, we have $2\gamma \leq \alpha + \gamma + \kappa < 2(\alpha + \gamma) \leq \nu$ so exactly as in (79) we have

$$(95) \quad S(\Delta^*, \ell^{\alpha+\gamma+\kappa}) = \begin{cases} \ell^{\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} & \text{if } \Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 0, 1 \pmod{4} \\ 0 & \text{if } \Delta/\ell^{2\lfloor \frac{\alpha+\gamma+\kappa}{2} \rfloor} \equiv 2, 3 \pmod{4}. \end{cases}$$

If $\Delta/\ell^{2\lfloor\frac{\alpha+\gamma+\kappa}{2}\rfloor} \equiv 0, 1 \pmod{4}$ then we have

$$(96) \quad c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \sum_{\substack{m^2|\Delta \\ v_\ell(m)=\kappa}} h_w \left(\frac{\Delta}{m^2} \right) \ell^{\lfloor\frac{\alpha+\gamma+\kappa}{2}\rfloor},$$

and again $\nu \geq 2(\alpha + \gamma)$ and $\kappa < \alpha + \gamma$ give $\nu - (\alpha + \gamma + \kappa) > 0$. Therefore, if $\Delta/\ell^{2\lfloor\frac{\alpha+\gamma+\kappa}{2}\rfloor} \equiv 0, 1 \pmod{4}$, then we may apply Lemma 11 and pair the terms exactly as in (81), (82), (83), and (84). In this case we have

$$\text{II} = \sum_{\kappa=\gamma-\alpha}^{\gamma+\alpha-1} c_\kappa(t, q, d) = D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) \sum_{j \leq \alpha-1} \sum_{\substack{m^2 | (\Delta/\ell^{2\gamma}) \\ v_\ell(m)=j}} h_w \left(\frac{\Delta}{\ell^{2\gamma} m^2} \right).$$

Therefore by the definition of H from (28), in the case ($\beta = \gamma$ and $\nu \geq 2(\alpha + \gamma)$) we have

$$\text{II} + \text{III} = D(t; \ell^{2\gamma}) \psi(\ell^{\alpha+\gamma}) H \left(\frac{\Delta}{\ell^{2\gamma}} \right).$$

The case $\kappa < \gamma - \alpha$, i.e. the sum I.

When we treated the case ($\beta = \gamma$, $\nu < 2(\alpha + \gamma)$ and $\kappa < \gamma - \alpha$) we did not use the assumption $\nu < 2(\alpha + \gamma)$ at all. Therefore the same proof goes through verbatim in the present case ($\beta = \gamma$, $\nu \geq 2(\alpha + \gamma)$ and $\kappa < \gamma - \alpha$). We take $C(t, q, d) = \text{I} + \text{II} + \text{III}$ and conclude that (90) holds whenever $\beta = \gamma$. Thus, by Lemma 12 we conclude Proposition 4 in all cases.

4.6. The General Case of Proposition 4. We now discuss the computation of $T^{(e)}(t)$ in the case of general levels M and N without giving full details. Recall that $(N, q) = 1$, $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, and $L = (d^2q - 1, N)$. Let $\ell \mid MN$ be a prime, $\alpha_\ell = v_\ell(M)$, $\beta_\ell = v_\ell(L)$ and $\gamma_\ell = v_\ell(N)$. Let

$$W(d) \stackrel{\text{def}}{=} \sum_{c \in (\mathbb{Z}/MN\mathbb{Z})^\times}^{*m} \delta_N(c, d^{-1})$$

and for positive integers K such that $K^2 \mid \Delta$, let

$$(97) \quad C_{K,N,M}(t, q, d) \stackrel{\text{def}}{=} \sum_{m^2 | (\Delta/K^2)} h_w \left(\frac{\Delta}{K^2 m^2} \right) \frac{\psi(MN)}{\psi(MN/(MN, m))} W(d).$$

Note that with this definition we have

$$C_{K,1,1}(t, q, d) = H \left(\frac{\Delta}{K^2} \right)$$

and

$$C_{1,\ell\gamma,\ell^\alpha}(t, q, d) = C(t, q, d).$$

Following the same steps as (48) through (50) we find that

$$T^{(e)}(t) = \frac{1}{2} (C_{1,N,M}(t, q, d) + (-1)^k C_{1,N,M}(t, q, -d)),$$

so it suffices to compute $C_{1,N,M}(t, q, d)$.

Proposition 5. *We have that*

$$\begin{aligned} & C_{K,N,M}(t, q, d) \\ &= \frac{\psi(\ell^{2\gamma_\ell})}{\psi(\ell^{2(\gamma_\ell - \alpha_\ell)})} \sum_{j=0}^{\beta_\ell - \alpha_\ell} \frac{\varphi(\ell^{2j})\varphi(\ell^{\gamma_\ell - \alpha_\ell - j})}{\varphi(\ell^{\gamma_\ell - \alpha_\ell})} \left(C_{\ell^{\alpha_\ell}K, N/\ell^{\gamma_\ell}, M/\ell^{\alpha_\ell}}(t, q, d) \delta_{\ell^{\alpha_\ell}}(d^2q, 1) D(t; \ell^{\alpha_\ell + \gamma_\ell}) \right. \\ &\quad \left. - \sum_{k=1}^{\gamma_\ell - \alpha_\ell - 1} C_{\ell^{\alpha_\ell + k}K, N/\ell^{\gamma_\ell}, M/\ell^{\alpha_\ell}}(t, q, d) \delta_{\ell^{\alpha_\ell + k}}(d^2q, 1) (D(t; \ell^{\alpha_\ell + \gamma_\ell + k - 1}) - D(t; \ell^{\alpha_\ell + \gamma_\ell + k})) \right). \end{aligned}$$

Proposition 4 is the special case of Proposition 5 with $K = 1$, $N = \ell^\gamma$ and $M = \ell^\alpha$. The proof of Proposition 5 is nearly the same as the proof of Proposition 4 but notationally more cumbersome. It suffices to replace within the proof of Proposition 4 the several instances of $H(\Delta/K^2)$ with $C_{K,N/\ell^{\gamma_\ell}, M/\ell^{\alpha_\ell}}(t, q, d)$ to pass to the proof of Proposition 5.

Applying Proposition 5 recursively, one checks that the result matches the definition of the class numbers $H_{N,M}(t, q, d)$ in (32). Therefore

$$(98) \quad C_{1,N,M}(t, q, d) = 2 \frac{\psi(N^2)}{\psi(N^2/M^2)} \sum_{\Lambda|(L/M)} \frac{\varphi(\Lambda^2)\varphi(N/(M\Lambda))}{\varphi(N/M)} H_{N,\Lambda M}(t, q, d)$$

and

$$T^{(e)}(t) = \frac{\psi(N^2)}{\psi(N^2/M^2)} \sum_{\Lambda|(L/M)} \frac{\varphi(\Lambda^2)\varphi(N/(M\Lambda))}{\varphi(N/M)} (H_{N,\Lambda M}(t, q, d) + (-1)^k H_{N,\Lambda M}(t, q, -d)).$$

Note from the definition that $H_{N,M}(t, q, -d) = H_{N,M}(-t, q, d)$ and also that $U_{k-2}(t, q)$ is an even (resp. odd) function of t when k is even (resp. odd). Therefore

$$(99) \quad \frac{1}{2} \sum_{t^2 < 4q} U_{k-2}(t, q) (H_{N,\Lambda M}(t, q, d) + (-1)^k H_{N,\Lambda M}(t, q, -d)) = \sum_{t^2 < 4q} U_{k-2}(t, q) H_{N,\Lambda M}(t, q, d)$$

and

$$T^{(e)} = \frac{\psi(N^2)}{\psi(N^2/M^2)} \sum_{\Lambda|(L/M)} \frac{\varphi(\Lambda^2)\varphi(N/(M\Lambda))}{\varphi(N/M)} \sum_{t^2 < 4q} U_{k-2}(t, q) H_{N,\Lambda M}(t, q, d),$$

as claimed in the statement of Theorem 9.

5. COMPARISON OF CLASS NUMBER SUMS

Recall the definitions of $\omega_A(q, d)$ and $\omega_A^*(q, d)$ from (40) and (41). The following expression was the main result of Section 3, expressing elliptic curve counts in terms of class numbers. See Proposition 3. We have

$$(100) \quad \begin{aligned} \mathbb{E}_q(U_{k-2}(t, q)\Phi_A) &= \frac{1}{q}\omega_A(q, d) + \frac{1}{q}\omega_A^*(q, 1) \\ &\quad + q^{k/2-1} \frac{(p-1)(k-1)}{24q} (\delta_{n_1(A)}(\sqrt{q}, 1) + (-1)^k \delta_{n_1(A)}(\sqrt{q}, -1)). \end{aligned}$$

The main result of Section 4, Theorem 9, expresses $\text{Tr}(T_q\langle d \rangle | S_k(n_1, n_2))$ as a sum of similar class numbers. Specifically, for $n_2 \mid n_1$ let

$$(101) \quad \Sigma_{n_1, n_2}(q, d) \stackrel{\text{def}}{=} \sum_{t^2 < 4q} U_{k-2}(t, q) H_{n_1, n_2}(t, q, d),$$

and if $q = p^{-1}$ we set $\Sigma_{n_1, n_2}(q, d) = 0$. Recall the definition of $\phi(n)$ as the Dirichlet convolution inverse to $\varphi(n^2)$ and the definition of $T_{n_1, \lambda}(q, d)$ in terms of traces of Hecke operators. Suppose $d^2 q \equiv 1 \pmod{n_2}$. The trace formula, Theorem 9, and the definition of ϕ imply that

$$(102) \quad \Sigma_{n_1, n_2}(q, d) = \frac{1}{\varphi(n_1/n_2)} \sum_{\nu \mid \frac{(d^2 q - 1, n_1)}{n_2}} \phi(\nu) T_{n_1, n_2 \nu}(q, d).$$

The goal of this section is the comparison of the expressions (100) and (102) given by Proposition 6 below. The main result of the paper, Theorem 3, follows directly from (100), (102), and Proposition 6.

Proposition 6. *Let A a finite abelian group of rank at most 2 and q be a power of a prime p such that $(q, |A|) = 1$. Then we have*

$$\omega_A(q, 1) + \omega_A^*(q, 1) = \Sigma_{n_1(A), n_2(A)}(q, 1) - p^{k-1} \Sigma_{n_1(A), n_2(A)}(q/p^2, p).$$

Proof. To lessen the notational burden, within this proof we write $U(t, q) = U_{k-2}(t, q)$ for the Chebyshev polynomials, $n_1 = n_1(A)$ and $n_2 = n_2(A)$, and take the d implicit in the $D(t; n)$ notation to be $d = 1$. Let $q = p^v$ with $v \in \mathbb{Z}_{\geq 0}$. To prove Proposition 6 it suffices to show that

$$(103) \quad \begin{aligned} \Sigma_{n_1, n_2}(q, 1) &= \sum_{0 \leq 2i < v} (p^i)^{k-1} (\omega_A(q/p^{2i}, p^i) + \omega_A^*(q/p^{2i}, p^i)) \\ &\quad + \frac{1}{2} p^{v/2} \delta(n_2, 1) \delta_2(v, 0) \left(\frac{1}{2} U(0, q) D(0; n_1) + \frac{1}{3} U(q^{1/2}, q) D(q^{1/2}, n_1) \right. \\ &\quad \left. + \frac{1}{3} U(-q^{1/2}, q) D(-q^{1/2}, n_1) \right) \end{aligned}$$

and

$$(104) \quad \begin{aligned} p^{k-1} \Sigma_{n_1, n_2}(q/p^2, p) &= \sum_{0 < 2i < v} (p^i)^{k-1} (\omega_A(q/p^{2i}, p^i) + \omega_A^*(q/p^{2i}, p^i)) \\ &\quad + \frac{1}{2} p^{v/2} \delta(n_2, 1) \delta_2(v, 0) \left(\frac{1}{2} U(0, q) D(0; n_1) + \frac{1}{3} U(q^{1/2}, q) D(q^{1/2}, n_1) \right. \\ &\quad \left. + \frac{1}{3} U(-q^{1/2}, q) D(-q^{1/2}, n_1) \right). \end{aligned}$$

Lemma 16. *Let $q = p^v$. For any i satisfying $0 \leq 2i < v$, we have that*

$$(105) \quad (p^i)^{k-1} \omega_A(q/p^{2i}, p^i d) = \sum_{\substack{v_p(t)=i \\ t^2 < 4q}} U(t, q) H_{n_1, n_2}(t, q, d).$$

Proof. Rearranging terms in the definition of $U_{k-2}(t, q)$ shows that

$$(106) \quad (p^i)^{k-1} U_{k-2}(t, q/p^{2i}) = p^i U_{k-2}(p^i t, q).$$

We claim that if $v_p(q/p^{2i}) \geq 1$ and $p \nmid t$ then

$$(107) \quad p^i H_{n_1, n_2}(t, q/p^{2i}, p^i) = H_{n_1, n_2}(p^i t, q, 1).$$

The lemma follows immediately after the change of variables $p^i t \rightarrow t$.

We now prove the claim. For any $(n, q) = 1$ such that $n^2 \mid ((p^i t)^2 - 4q)$ we have

$$\begin{aligned} H\left(\frac{(p^i t)^2 - 4q}{n^2}\right) &= \sum_{\delta^2 \mid \frac{(p^i t)^2 - 4q}{n^2}} h_w\left(\frac{(p^i t)^2 - 4q}{n^2 \delta^2}\right) \\ &= \sum_{\kappa=0}^i \sum_{\substack{\delta^2 \mid \frac{(p^{i-\kappa} t)^2 - 4q/p^{2\kappa}}{n^2} \\ v_p(\delta)=0}} h_w\left(\frac{(p^{i-\kappa} t)^2 - 4q/p^{2\kappa}}{n^2 \delta^2}\right). \end{aligned}$$

For any κ satisfying $0 \leq \kappa \leq i-1$ we apply Lemma 11 with $d = \frac{t^2 - 4q/p^{2i}}{n^2}$ and $f = p^{i-\kappa}$. Since $(n, p) = 1$ and $p \nmid t$ we have that $\frac{t^2 - 4q/p^{2i}}{n^2}$ is a non-zero square modulo p . This implies

$$\begin{aligned} H\left(\frac{(p^i t)^2 - 4q}{n^2}\right) &= \sum_{\substack{\delta^2 \mid \frac{t^2 - 4q/p^{2i}}{n^2} \\ v_p(\delta)=0}} h_w\left(\frac{t^2 - 4q/p^{2i}}{n^2 \delta^2}\right) \\ &\quad + \sum_{\kappa=0}^{i-1} p^{i-\kappa} \left(1 - \frac{1}{p}\right) \sum_{\substack{\delta^2 \mid \frac{t^2 - 4q/p^{2i}}{n^2} \\ v_p(\delta)=0}} h_w\left(\frac{t^2 - 4q/p^{2i}}{n^2 \delta^2}\right) \\ &= p^i H\left(\frac{t^2 - 4q/p^{2i}}{n^2}\right). \end{aligned}$$

We also have that $n \mid (p^{-i} dq - p^{-i} d^{-1} - t)$ if and only if $n \mid (dq - d^{-1} - p^i t)$ since $(n, p) = 1$. This completes the proof of the claim. \square

Verifying (103).

We prove that equation (103) holds by considering several cases. Recall that $q = p^v$.

If $q = p^v$ with v even, then taking the sum of (105) over i gives

$$(108) \quad \begin{aligned} \Sigma_{n_1, n_2}(q, 1) &= \sum_{0 \leq 2i < v} (p^i)^{k-1} \omega_A(q/p^{2i}, p^i) + U(-q^{1/2}, q) H_{n_1, n_2}(-q^{1/2}, q, 1) \\ &\quad + U(0, q) H_{n_1, n_2}(0, q, 1) + U(q^{1/2}, q) H_{n_1, n_2}(q^{1/2}, q, 1). \end{aligned}$$

If $q = p^v$ with v odd, then taking the sum of (105) over i gives

$$(109) \quad \Sigma_{n_1, n_2}(q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} \omega_A(q/p^{2i}, p^i) + U(0, q) H_{n_1, n_2}(0, q, 1) \\ + \begin{cases} 0 & \text{if } p \neq 2, 3 \\ U(-2^{\frac{v+1}{2}}, q) H_{n_1, n_2}(-2^{\frac{v+1}{2}}, q, 1) + U(2^{\frac{v+1}{2}}, q) H_{n_1, n_2}(2^{\frac{v+1}{2}}, q, 1) & \text{if } p = 2 \\ U(-3^{\frac{v+1}{2}}, q) H_{n_1, n_2}(-3^{\frac{v+1}{2}}, q, 1) + U(3^{\frac{v+1}{2}}, q) H_{n_1, n_2}(3^{\frac{v+1}{2}}, q, 1) & \text{if } p = 3. \end{cases}$$

Similarly, it follows from Lemma 16 when v is even that

$$(110) \quad p^{k-1} \Sigma_{n_1, n_2}(q/p^2, p) = \sum_{0 < 2i < v} (p^i)^{k-1} \omega_A(q/p^{2i}, p^i) \\ + p^{k-1} (U(-q^{1/2}/p, q/p^2) H_{n_1, n_2}(-q^{1/2}/p, q/p^2, p) + U(0, q/p^2) H_{n_1, n_2}(0, q/p^2, p) \\ + U(q^{1/2}/p, q/p^2) H_{n_1, n_2}(q^{1/2}/p, q/p^2, p)),$$

and when v is odd that

$$(111) \quad p^{k-1} \Sigma_{n_1, n_2}(q/p^2, p) = \sum_{0 < 2i < v} (p^i)^{k-1} \omega_A(q/p^{2i}, p^i) + p^{k-1} U(0, q/p^2) H_{n_1, n_2}(0, q/p^2, p) \\ + p^{k-1} \begin{cases} 0 & \text{if } p \neq 2, 3 \\ U(-2^{\frac{v-1}{2}}, 2^{v-2}) H_{n_1, n_2}(-2^{\frac{v-1}{2}}, 2^{v-2}, 2) \\ \quad + U(2^{\frac{v-1}{2}}, 2^{v-2}) H_{n_1, n_2}(2^{\frac{v-1}{2}}, 2^{v-2}, 2) & \text{if } p = 2 \\ U(-3^{\frac{v-1}{2}}, 3^{v-2}) H_{n_1, n_2}(-3^{\frac{v-1}{2}}, 3^{v-2}, 3) \\ \quad + U(3^{\frac{v-1}{2}}, 3^{v-2}) H_{n_1, n_2}(3^{\frac{v-1}{2}}, 3^{v-2}, 3) & \text{if } p = 3. \end{cases}$$

Returning to (108) and (109) we calculate from the definition of $H_{n_1, n_2}(t, q, d)$ that:

(1) We have

$$(112) \quad H_{n_1, n_2}(0, q, 1) = \begin{cases} 0 & \text{if } n_2 > 2 \\ \frac{1}{2} H(-q) D(0, 2n_1) & \text{if } n_2 = 2 \\ \frac{1}{2} H(-4q) D(0, n_1) \\ \quad - \delta_4(n_1, 0) \frac{1}{2} H(-q) (D(0, n_1) - D(0, 2n_1)) & \text{if } n_2 = 1. \end{cases}$$

(2) If v even, then

$$(113) \quad H_{n_1, n_2}(\pm q^{1/2}, q, 1) = \begin{cases} 0 & \text{if } n_2 > 1 \\ \frac{1}{2} H(-3q) D(\pm q^{1/2}, n_1) & \text{if } n_2 = 1. \end{cases}$$

(3) If v odd and $p = 2$, then

$$(114) \quad H_{n_1, n_2}(\pm 2^{\frac{v+1}{2}}, q, 1) = \begin{cases} 0 & \text{if } n_2 > 1 \\ \frac{1}{2} H(-2q) D(\pm 2^{\frac{v+1}{2}}, n_1) & \text{if } n_2 = 1. \end{cases}$$

(4) If v odd and $p = 3$, then

$$(115) \quad H_{n_1, n_2}(\pm 3^{\frac{v+1}{2}}, q, 1) = \begin{cases} 0 & \text{if } n_2 > 1 \\ \frac{1}{2}H(-q)D(\pm 3^{\frac{v+1}{2}}, n_1) & \text{if } n_2 = 1. \end{cases}$$

We check that (103) is true by separating into cases $n_2 > 2$, $n_2 = 2$, and $n_2 = 1$.

The case $n_2 > 2$.

If $n_2 > 2$ then (112), (113), (114), and (115) all vanish, but so does $\omega_A^*(q, 1)$ by definition of $H_{n_1, n_2}^*(t, q, d)$. Therefore we have verified (103) in the case that $n_2 > 2$.

The case $n_2 = 2$.

If $n_2 = 2$ then of (112), (113), (114), and (115) only $H_{n_1, n_2}(0, q, 1) = \frac{1}{2}H(-q)\delta_{2n_1}(q+1, 0)$ can be non-zero. If $\delta_{2n_1}(q+1, 0) = 1$, then $q \equiv 3 \pmod{4}$, and v must be odd. Then Lemma 11, (106), and $(n, p) = 1$ imply that

$$\begin{aligned} U(0, q)H_{n_1, n_2}(0, q, 1) &= \frac{1}{2}h_w(-p)\sigma(p^{\frac{v-1}{2}})U(0, q)\delta_{2n_1}(q+1, 0) \\ &= \frac{1}{2}h_w(-p) \sum_{0 \leq 2i < v} (p^i)^{k-1} U(0, q/p^{2i})\delta_{2n_1}(p^i q/p^{2i} + p^{-i}, 0). \end{aligned}$$

The definition (37) of H_{n_1, n_2}^* and the definition of ω_A^* imply that

$$(116) \quad \begin{aligned} U(0, q)H_{n_1, n_2}(0, q, 1) &= \sum_{0 \leq 2i < v} (p^i)^{k-1} \sum_{t^2 < 4q} U(t, q/p^{2i})H_{n_1, n_2}^*(t, q/p^{2i}, p^i) \\ &= \sum_{0 \leq 2i < v} (p^i)^{k-1} \omega_A^*(q/p^{2i}, p^i). \end{aligned}$$

Therefore we have verified (103) in the case that $n_2 = 2$.

The case $n_2 = 1$ and v even.

In this case, $q \equiv 1 \pmod{4}$ and we have by Lemma 11 and (112) that

$$H_{n_1, 1}(0, q, 1) = \frac{1}{4} \left(\sigma(p^{v/2-1}) \left(1 - \left(\frac{-4}{p} \right) \right) + p^{v/2} \right) D(0, n_1).$$

This expression together with (38) implies that

$$(117) \quad U(0, q)H_{n_1, 1}(0, q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} U(0, q/p^{2i})H_{n_1, 1}^*(0, q/p^{2i}, p^i) + \frac{1}{4}p^{v/2}U(0, q)D(0, n_1).$$

Likewise, if $t = \pm q^{1/2}$ and v is even, we apply Lemma 11 to (113) to find that

$$H_{n_1, 1}(\pm q^{1/2}, q, 1) = \frac{1}{6} \left(\sigma(p^{v/2-1}) \left(1 - \left(\frac{-3}{p} \right) \right) + p^{v/2} \right) D(\pm q^{1/2}, n_1).$$

By (38) we have

$$(118) \quad U(\pm q^{1/2}, q)H_{n_1,1}(\pm q^{1/2}, q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} U(\pm q^{1/2}, q/p^{2i}) H_{n_1,1}^*(\pm q^{1/2}, q/p^{2i}, p^i) \\ + \frac{1}{6} p^{v/2} U(\pm q^{1/2}, q) D(\pm q^{1/2}, n_1).$$

Putting together (108), (117), and (118) we verify (103) in the case that v is even and $n_2 = 1$.

The case $n_2 = 1$ and v odd.

In this case, we have by (112) and Lemma 11 that

$$H_{n_1,1}(0, q, 1) = \frac{1}{2} \sigma(p^{\frac{v-1}{2}}) (H(-4p)D(0, n_1) - \delta_4(n_1, 0)h_w(-p) (D(0, n_1) - D(0, 2n_1))).$$

It follows from (39) and (106) that

$$(119) \quad U(0, q)H_{n_1,1}(0, q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} U(0, q/p^{2i}) H_{n_1,1}^*(0, q/p^{2i}, p^i).$$

If $n_2 = 1$, v is odd, and $p = 2$ then by (114) and Lemma 11, we have that

$$H_{n_1, n_2}(\pm 2^{\frac{v+1}{2}}, q, 1) = \frac{1}{4} \sigma(2^{\frac{v-1}{2}}) D(\pm 2^{\frac{v+1}{2}}, n_1).$$

This expression together with (39) and (106) imply that

$$(120) \quad U(\pm 2^{\frac{v+1}{2}}, q)H_{n_1, n_2}(\pm 2^{\frac{v+1}{2}}, q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} U(\pm 2^{\frac{v+1}{2}}, q/p^{2i}) H_{n_1,1}^*(\pm 2^{\frac{v+1}{2}}, q/p^{2i}, p^i).$$

Similarly, if $n_2 = 1$, v is odd, and $p = 3$ we have that

$$(121) \quad U(\pm 3^{\frac{v+1}{2}}, q)H_{n_1, n_2}(\pm 3^{\frac{v+1}{2}}, q, 1) = \sum_{0 \leq 2i < v} (p^i)^{k-1} U(\pm 3^{\frac{v+1}{2}}, q/p^{2i}) H_{n_1,1}^*(\pm 3^{\frac{v+1}{2}}, q/p^{2i}, p^i).$$

Putting together (109) and (119) we verify (103) in the case that $p \neq 2, 3$, v is odd, and $n_2 = 1$. If $p = 2$, v is odd, and $n_2 = 1$ we apply (109), (119), and (120) to verify (103). In the case $p = 3$, v is odd, and $n_2 = 1$ we apply (109), (119), and (121) to verify (103). We have now verified (103) in all cases.

Verifying (104).

To verify this equation we return to (110) and (111) and perform similar calculations to those given above.

The case $n_2 > 2$.

In this case, each of the terms (112), (113), (114), and (115) vanish, but so do $H_{n_1, n_2}^*(t, q/p^2, p)$ and $\omega_A^*(q/p^2, p)$, so we have verified (104) in this case.

The case $n_2 = 2$.

In this case, by (106) and (112)

$$p^{k-1}U(0, q/p^2)H_{n_1, n_2}(0, q/p^2, p) = pU(0, q)\frac{1}{2}H(-q/p^2)D(0, 2n_1)$$

and by Lemma 11 we have

$$p\frac{1}{2}H(-q/p^2) = \frac{1}{2}p\sigma(p^{\frac{v-3}{2}})h_w(-p) = \frac{1}{2}\left(\sigma(p^{\frac{v-1}{2}}) - 1\right)h_w(-p) = \frac{1}{2}H(-q) - \frac{1}{2}h_w(-p).$$

A similar argument to the one that gave (116) implies that

$$\begin{aligned} & pU(0, q)H(0, q/p^2, p) \\ &= U(0, q)\frac{1}{2}H(-q)D(0, 2n_1) - U(0, q)\frac{1}{2}h_q(-p)D(0, 2n_1) \\ &= \sum_{0 \leq 2i < v} (p^i)^{k-1} \sum_{t^2 < 4q} U(t, q/p^{2i})H_{n_1, 2}^*(t, q/p^{2i}, p^i) - \frac{1}{2}h_w(-p)U(0, q)D(0, 2n_1) \\ &= \sum_{0 < 2i < v} (p^i)^{k-1} \sum_{t^2 < 4q} U(t, q/p^{2i})H_{n_1, 2}^*(t, q/p^{2i}, p^i), \end{aligned}$$

which verifies (104) in the case $n_2 = 2$.

The case $n_2 = 1$.

By (112) and (106) we have

$$\begin{aligned} p^{k-1}U(0, q/p^2)H_{n_1, 1}(0, q/p^2, p) &= pU(0, q)\left(\frac{1}{2}H(-4q/p^2)D(0, n_1) \right. \\ &\quad \left. - \delta_4(n_1, 0)\frac{1}{2}H(-q/p^2)(D(0, n_1) - D(0, 2n_1))\right). \end{aligned}$$

Lemma 11 implies that

$$p\frac{1}{2}H(-q/p^2) = \frac{1}{2}H(-q) - \frac{1}{2}h_w(-p)$$

and

$$p\frac{1}{2}H(-4q/p^2) = \frac{1}{2}H(-4q) - \begin{cases} \frac{1}{4}\left(1 - \left(\frac{-4}{p}\right)\right) & \text{if } v \text{ even} \\ \frac{1}{2}H(-4p) & \text{if } v \text{ odd.} \end{cases}$$

Therefore, by (117) and (119) we have that

$$\begin{aligned} p^{k-1}U(0, q/p^2)H_{n_1, 1}(0, q/p^2, p) &= \sum_{0 \leq 2i < v} (p^i)^{k-1}U(0, q/p^{2i})H_{n_1, 1}^*(0, q/p^{2i}, p^i) \\ &\quad + \frac{1}{4}p^{v/2}U(0, q)D(0, n_1) - U(0, q)D(0, n_1) \begin{cases} \frac{1}{4}\left(1 - \left(\frac{-4}{p}\right)\right) & \text{if } v \text{ even} \\ \frac{1}{2}H(-4p) & \text{if } v \text{ odd} \end{cases} \\ &\quad + \delta_4(n_1, 0)\frac{1}{2}h_w(-p)(D(0, n_1) - D(0, 2n_1)). \end{aligned}$$

Applying (38) or (39) to this expression gives

$$(122) \quad p^{k-1}U(0, q/p^2)H_{n_1,1}(0, q/p^2, p) = \sum_{0 < 2i < v} (p^i)^{k-1}U(0, q/p^{2i})H_{n_1,1}^*(0, q/p^{2i}, p^i) \\ + \frac{1}{4}p^{v/2}U(0, q)D(0, n_1).$$

The case $n_2 = 1$ and v even.

By (106), (113), and $(n_1, p) = 1$ we have

$$p^{k-1}U(\pm q^{1/2}/p)H_{n_1, n_2}(\pm q^{1/2}/q, q/p^2, p) = pU(\pm q^{1/2}, q)\frac{1}{2}H(-3q/p^2)D(\pm q^{1/2}, n_1).$$

By Lemma 11 we have

$$\begin{aligned} p\frac{1}{2}H(-3q/p^2) &= \frac{1}{6} \left(p\sigma(p^{\frac{v}{2}-2}) \left(1 - \left(\frac{-3}{p} \right) \right) + p^{v/2} \right) \\ &= \frac{1}{6} \left((\sigma(p^{\frac{v}{2}-1}) - 1) \left(1 - \left(\frac{-3}{p} \right) \right) + p^{v/2} \right) \\ &= \frac{1}{2}H(-3q) - \frac{1}{6} \left(1 - \left(\frac{-3}{p} \right) \right). \end{aligned}$$

Using these last two equations we find by (118) that

$$\begin{aligned} p^{k-1}U(\pm q^{1/2}/p)H_{n_1, n_2}(\pm q^{1/2}/q, q/p^2, p) \\ = \sum_{0 \leq 2i < v} (p^i)^{k-1}U(\pm q^{1/2}, q/p^{2i})H_{n_1,1}^*(\pm q^{1/2}, q/p^{2i}, p^i) + \frac{1}{6}p^{v/2}U(\pm q^{1/2}, q)D(\pm q^{1/2}, n_1) \\ - U(\pm q^{1/2}, q)D(\pm q^{1/2}, n_1)\frac{1}{6} \left(1 - \left(\frac{-3}{p} \right) \right). \end{aligned}$$

By the definition (38) of $H_{n_1,1}^*$ this simplifies to

$$(123) \quad p^{k-1}U(\pm q^{1/2}/p)H_{n_1, n_2}(\pm q^{1/2}/q, q/p^2, p) \\ = \sum_{0 < 2i < v} (p^i)^{k-1}U(\pm q^{1/2}, q/p^{2i})H_{n_1,1}^*(\pm q^{1/2}, q/p^{2i}, p^i) + \frac{1}{6}p^{v/2}U(\pm q^{1/2}, q)D(\pm q^{1/2}, n_1).$$

Putting together (111), (122), and (123) verifies (104) when v is even.

The case $n_2 = 1$ and v odd.

If v is odd and $p = 2$ then by $(n_1, 2) = 1$, (106), and (114) we have

$$2^{k-1}U(\pm 2^{\frac{v-1}{2}}, 2^{v-2})H_{n_1,1}(\pm 2^{\frac{v-1}{2}}, 2^{v-2}, 2) = 2U(\pm 2^{\frac{v+1}{2}}, 2^v)\frac{1}{2}H(-2q/2^2)D(\pm 2^{\frac{v+1}{2}}, n_1).$$

We have by Lemma 11 that

$$\frac{1}{2}H(-2q/2^2) = \frac{1}{4} \left(\sigma(2^{\frac{v-1}{2}}) - 1 \right) = \frac{1}{2}H(-2q) - \frac{1}{4}.$$

Therefore by (120) we have

$$\begin{aligned} & 2^{k-1}U(\pm 2^{\frac{v-1}{2}}, 2^{v-2})H_{n_1,1}(\pm 2^{\frac{v-1}{2}}, 2^{v-2}, 2) \\ &= \sum_{0 \leq 2i < v} (p^i)^{k-1}U(\pm 2^{\frac{v+1}{2}}, q/p^{2i})H_{n_1,1}^*(\pm 2^{\frac{v+1}{2}}, q/p^{2i}, p^i) - \frac{1}{4}U(\pm 2^{\frac{v+1}{2}}, 2^v)D(2^{\frac{v+1}{2}}, n_1). \end{aligned}$$

By the definition (39) of $H_{n_1,1}^*(t, q, d)$ this simplifies to

$$\begin{aligned} (124) \quad & 2^{k-1}U(\pm 2^{\frac{v-1}{2}}, 2^{v-2})H_{n_1,1}(\pm 2^{\frac{v-1}{2}}, 2^{v-2}, 2) \\ &= \sum_{0 < 2i < v} (p^i)^{k-1}U(\pm 2^{\frac{v+1}{2}}, q/p^{2i})H_{n_1,1}^*(\pm 2^{\frac{v+1}{2}}, q/p^{2i}, p^i). \end{aligned}$$

Similarly, if v is odd and $p = 3$ we have by (114) and (121) that

$$\begin{aligned} (125) \quad & 3^{k-1}U(\pm 3^{\frac{v-1}{2}}, 3^{v-2})H_{n_1,1}(\pm 3^{\frac{v-1}{2}}, 3^{v-2}, 3) \\ &= \sum_{0 < 2i < v} (p^i)^{k-1}U(\pm 3^{\frac{v+1}{2}}, q/p^{2i})H_{n_1,1}^*(\pm 3^{\frac{v+1}{2}}, q/p^{2i}, p^i). \end{aligned}$$

Combining (111), (122), (124), and (125) verifies (104) when v is odd. \square

6. ACKNOWLEDGEMENTS

The authors thank Alina Cojocaru, Noam Elkies, Jordan Ellenberg, and Corentin Perret-Gentil for helpful conversations related to this project, and also the anonymous referee whose detailed report greatly improved this paper.

REFERENCES

- [1] R. Bell, C. Blakestad, A. C. Cojocaru, A. Cowan, N. Jones, V. Matei, G. Smith, and I. Vogt. Constants in Titchmarsh divisor problems for elliptic curves. Preprint, arXiv:1706.03422, 2017.
- [2] B. J. Birch. How the number of points of an elliptic curve over a fixed prime field varies. *J. London Math. Soc.*, 43:57–60, 1968.
- [3] Bradley W. Brock and Andrew Granville. More points than expected on curves over finite field extensions. *Finite Fields Appl.*, 7(1):70–91, 2001.
- [4] Henri Cohen. Trace des opérateurs de Hecke sur $\Gamma_0(N)$. *Séminaire de Théorie des Nombres (1976–1977)*, Exp. No. 4, 9 pp. *Lab. Théorie des Nombres, CNRS, Talence*, 1977.
- [5] A. C. Cojocaru. Averages of elliptic curves modulo p with a prescribed first elementary divisor. In Preparation, 2017.
- [6] J. B. Conrey, W. Duke, and D. W. Farmer. The distribution of the eigenvalues of Hecke operators. *Acta Arith.*, 78(4):405–409, 1997.
- [7] David A. Cox. *Primes of the form $x^2 + ny^2$, Fermat, class field theory, and complex multiplication*. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2013.
- [8] Max Deuring. Die Typen der Multiplikatorenringe elliptischer Funktionenkörper. *Abh. Math. Sem. Hansischen Univ.*, 14:197–272, 1941.
- [9] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.
- [10] Ernst-Ulrich Gekeler. The distribution of group structures on elliptic curves over finite prime fields. *Doc. Math.*, 11:119–142 (electronic), 2006.
- [11] D. Roger Heath-Brown. The fourth power mean of Dirichlet’s L -functions. *Analysis* 1 (1981), no. 1, 25–32.
- [12] Hiroaki Hijikata. Explicit formula of the traces of Hecke operators for $\Gamma_0(N)$. *J. Math. Soc. Japan*, 26:56–82, 1974.

- [13] Everett W. Howe. On the group orders of elliptic curves over finite fields. *Compositio Math.*, 85(2):229–247, 1993.
- [14] J. Igusa, Fibre systems of Jacobian varieties III. Fibre systems of elliptic curves. *Amer. J. Math.* 81:453–476, 1959.
- [15] Y. Ihara, Hecke polynomials as congruence ζ functions in elliptic modular case. *Ann. of Math. (2)*, 85:267–295, 1967.
- [16] Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory*. American Mathematical Society Colloquium Publications, 2004.
- [17] Camille Jordan. Sur la forme canonique des congruences du second degré et le nombre de leurs solutions. *Journal de mathématiques pures et appliquées 2e série*, tome 17 (1872): 368–402.
- [18] Nathan Kaplan and Ian Petrow. Traces of Hecke operators and refined weight enumerators of Reed-Solomon codes. To appear in *Trans. Amer. Math. Soc.*, 2018, DOI: <https://doi.org/10.1090/tran/7089>
- [19] Nicholas Katz Frobenius-Schur indicator and the ubiquity of Brock-Granville quadratic excess. *Finite Fields Appl.*, 7(1):45–69, 2001.
- [20] Nicholas Katz and Peter C. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*. American Mathematical Society Colloquium Publications, 1999.
- [21] Andrew Knightly and Charles Li. Traces of Hecke operators (excerpt). <http://www.charlesli.org/math/papers/tracesexcerpt.pdf>. Accessed: 2015-09-21.
- [22] Andrew Knightly and Charles Li. *Traces of Hecke operators*, volume 133 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.
- [23] Paul Koosis. *The Logarithmic Integral*. Cambridge University Press, Cambridge, 1998.
- [24] Emmanuel Kowalski. Analytic problems for elliptic curves. *J. Ramanujan Math. Soc.* 21(1):19–114, 2006.
- [25] H. W. Lenstra Jr. Factoring integers with elliptic curves. *Ann. of Math.*, 2(3):649–673, 1987.
- [26] J. Oesterlé. Sur la trace des opérateurs de Hecke. Thèse, *L'Université de Paris-Sud Centre d'Orsay*, 1977.
- [27] Jordi Quer. Dimensions of spaces of modular forms for $\Gamma_H(N)$. *Acta Arith.*, 145(4):373–395, 2010.
- [28] Shepley L. Ross, II. A simplified trace formula for Hecke operators for $\Gamma_0(N)$. *Trans. Amer. Math. Soc.*, 331(1):425–447, 1992.
- [29] René Schoof. Nonsingular plane cubic curves over finite fields. *J. Combin. Theory Ser. A*, 46(2):183–211, 1987.
- [30] Atle Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc. (N.S.)*, 20:47–87, 1956.
- [31] Atle Selberg. On the estimation of Fourier coefficients of modular forms. Proceedings of Symposia on Pure Mathematics, Volume VIII, pages 1–15, 1965.
- [32] Jean-Pierre Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p . *J. Amer. Math. Soc.*, 10(1):75–102, 1997.
- [33] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*. Reprint of the 1971 original. Publications of the Mathematical Society of Japan, 11. Princeton University Press, Princeton, NJ, 1994.
- [34] S. G. Vlăduț. Cyclicity statistics for elliptic curves over finite fields. *Finite Fields Appl.*, 5(1):13–25, 1999.
- [35] W. Waterhouse. Abelian varieties over finite fields. *Ann. Sci. École Norm. Sup.*, 2: 521–560, 1969.

NATHAN KAPLAN – DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, IRVINE, CA 92697,
NCKAPLAN@MATH.UCI.EDU

IAN PETROW – ECOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SECTION DES MATHÉMATIQUES,
1015 LAUSANNE, SWITZERLAND, IAN.PETROW@EPFL.CH

Current address: ETH Zürich - Departement Mathematik HG G 66.4 Rämistrasse 101, 8092 Zürich,
Switzerland, ian.petrow@math.ethz.ch