AN AVERAGE ASYMPTOTIC FOR THE NUMBER OF EXTREMAL PRIMES OF ELLIPTIC CURVES

LUKE GIBERSON AND KEVIN JAMES

ABSTRACT. Let $E/\mathbb{Q}$ be an elliptic curve, and let p be a rational prime of good reduction. Let $a_p(E)$ denote the trace of the Frobenius endomorphism of E at the prime p. We say p is a *champion prime of* E if $a_p(E) = -\left[\left|2\sqrt{p}\right|\right]$, which occurs precisely when the group of $\mathbb{F}_p$-rational points is as large as possible in accordance with the Hasse bound. In a similar vein, we say p is a *trailing prime of* E if $a_p(E) = +\left[\left|2\sqrt{p}\right|\right]$, which occurs precisely when the group of $\mathbb{F}_p$-rational points is as small as possible in accordance with the Hasse bound. Together, we say that these primes constitute the *extremal primes of* E. In this paper we prove that on average, the number of champion primes of elliptic curves that are less than X is asymptotically equal to $\frac{8}{3\pi} \cdot X^{1/4}/\log X$. As an immediate corollary, we also gain asymptotics on the average number of trailing primes less than X and the average number of extremal primes less than X.

1. INTRODUCTION

For an elliptic curve $E/\mathbb{Q}$ and a prime p of good reduction, let $\bar{E}/\mathbb{F}_p$ denote the reduction of the curve E modulo p. In this setting, $\#\bar{E}/\mathbb{F}_p$ is the number of $\mathbb{F}_p$-rational points on $\bar{E}/\mathbb{F}_p$, and the trace of Frobenius of E at p is the integer $a_p(E)$ such that $\#\bar{E}/\mathbb{F}_p = p + 1 - a_p(E)$. A classical result of Hasse (see [Sil86, Theorem V.1.1]) is the bound $|a_p(E)| \leq 2\sqrt{p}$. In this way, we refer to the interval $[-2\sqrt{p}, 2\sqrt{p}]$ as the Hasse interval of E at p.

Of particular interest is a statistical investigation of the sequence $\{a_p(E)\}$ for a fixed curve E. For instance, we can normalize the traces $a_p(E)$ by defining the associated $b_p(E) := a_p(E)/2\sqrt{p}$. We have a strong understanding of the distribution of the normalized traces, as shown by the Sato-Tate conjecture (now a theorem – see [CHT08, Tay08, HSBT10, BLGHT11]).

Theorem 1.1 (Clozel, Harris, Shepherd-Barron, Taylor)**.** *Let $E/\mathbb{Q}$ be a curve without complex multiplication, and let $[a, b] \subset [-1, 1]$. Then as $X \to \infty$, we have*

$$\#\{p < X : b_p(E) \in [a, b]\} \quad \sim \quad \frac{2}{\pi}\left(\int_a^b \sqrt{1-t^2}\ \mathrm{d}t\right)\frac{X}{\log X}.$$

The distribution of fixed traces is not as well understood. For an elliptic curve $E/\mathbb{Q}$ and an integer $r \neq 0$, Lang and Trotter conjecture in [LT76]

2010 *Mathematics Subject Classification.* Primary 11G05; Secondary 11F30.

Key words and phrases. elliptic curves, extremal primes, champion primes, average Frobenius distributions.

26 Oct 2017 17:44:55 PDT
Version 2 - Submitted to Acta Arith.

that

$$\#\{p < X : a_p(E) = r\} \quad \sim \quad C_{E,r}\frac{\sqrt{X}}{\log X},$$

where $C_{E,r}$ is an explicitly defined constant depending only on E and r.

In this paper, we investigate the frequency at which the trace of Frobenius $a_p(E)$ is maximal or minimal inside the Hasse interval. This has a slightly different flavor than the Lang-Trotter conjecture mentioned above seeing that the "target" for the trace of Frobenius is changing with the prime p.

Let E be an elliptic curve, and let p be a rational prime of good reduction. We make the following definitions:

(1) p is a *champion prime of* E if $a_p(E) = -\left[\left|2\sqrt{p}\right|\right]$;
(2) p is a *trailing prime of* E if $a_p(E) = +\left[\left|2\sqrt{p}\right|\right]$;
(3) p is a *extremal prime of* E if $a_p(E) = \pm\left[\left|2\sqrt{p}\right|\right]$.

The study of extremal primes was initiated by Hedetniemi, James, and Xue in [HJX14], in which the authors consider primes such that $a_p(E) = -\left[\left|2\sqrt{p}\right|\right]$. They prove the following theorem, which establishes that almost all elliptic curves have at least one champion prime (and hence at least one extremal prime). In fact, the same method also establishes that almost all elliptic curves have at least one trailing prime.

Theorem 1.2 (Hedetniemi, James, Xue)**.** *Let X be a positive real number, and let $A := A(X), B := B(X)$ be positive parameters depending only on X. For any $\epsilon > 0$, take*

$$\begin{aligned} A, B &\geq \exp((1/4+\epsilon)X), \\ AB &\geq \exp((5/4+\epsilon)X). \end{aligned}$$

For any $a, b \in \mathbb{Z}$ such that $4a^3 + 27b^2 \neq 0$, let $E(a,b)$ be the elliptic curve given by the affine equation $y^2 = x^3 + ax + b$. Define sets

$$\begin{aligned} \mathcal{E}(A,B) &= \{E(a,b) : |a| \leq A, |b| \leq B\} \\ \mathcal{E}^-(A,B) &= \{E(a,b) \in \mathcal{E}(A,B) : E(a,b) \text{ has a champion prime}\}. \end{aligned}$$

Then $\#\mathcal{E}^-(A,B) \sim \#\mathcal{E}(A,B)$ as $X \to \infty$.

In [JTT+16], the authors established an asymptotic on the number of champion primes up to X for a elliptic curve $E/\mathbb{Q}$ with complex multiplication (CM), but this result was conditional on the assumption of the Riemann Hypothesis for certain Hecke L-functions. A recent paper of James and Pollack (see [JP17]) has removed this assumption.

Theorem 1.3 (James, Pollack)**.** *Let $E/\mathbb{Q}$ be an elliptic curve with complex multiplication. The number of champion primes $p < X$ is asymptotically $\frac{2X^{3/4}}{3\pi \log X}$. The number of trailing primes $p < X$ has an identical asymptotic.*

Apart from Theorem 1.2, nothing is known about extremal primes on non-CM curves. In this paper we will focus on counting champion primes on average over the family of all elliptic curves and mention in passing how

the results can be used to study trailing and extremal primes as well. For an elliptic curve $E/\mathbb{Q}$ and a positive real number X, we define

$$\pi_E^{\text{Champ}}(X) = \#\{p < X \text{ of good reduction of } E : a_p(E) = -\left[|2\sqrt{p}|\right]\}.$$

We note that because the Sato-Tate distribution is much different in the CM versus the non-CM case, the predicted asymptotic for the non-CM case is much smaller than the asymptotic obtained by James and Pollack in Theorem 1.3. Arguing heuristically with Sato-Tate, for a non-CM elliptic curve E, the "probability" that $a_p(E) = -\left[\left|2\sqrt{p}\right|\right]$ is approximately

$$\begin{aligned}\frac{2}{\pi}\int_{-1}^{-1+\frac{1}{2\sqrt{p}}}\sqrt{1-t^2}\,\mathrm{d}t &= \frac{2}{\pi}\int_{-1}^{-1+\frac{1}{2\sqrt{p}}}\left(\sqrt{2}(1-t)^{1/2}+\mathrm{O}((1-t)^{3/2}))\right)\,\mathrm{d}t\\ &= \frac{2}{\pi}\left(\frac{2\sqrt{2}}{3}\left(\frac{1}{2\sqrt{p}}\right)^{3/2}\right)+\mathrm{O}(p^{-5/4})\\ &= \frac{2}{3\pi}p^{-3/4}+\mathrm{O}(p^{-5/4}),\end{aligned}$$

where in the first equality we have exploited the identity

$$1-t^2 = 2(1-t)-(1-t)^2$$

and subsequently used a first order Taylor approximation on $\sqrt{1-t^2}$. Assuming independence and summing over all primes while ignoring error terms gives the expectation that

$$\pi_E^{\text{Champ}}(X) \sim \frac{2}{3\pi}\sum_{p<X}p^{-3/4} \sim \frac{8X^{1/4}}{3\pi\log X}.$$

The main result of this paper indicates that this heuristic is correct on average.

Theorem 1.4. *For every $A, B \geq 1$, we have*

$$\frac{1}{4AB}\sum_{\substack{|a|\leq A\\ |b|\leq B}}\pi_{E(a,b)}^{\text{Champ}}(X) = \frac{8X^{1/4}}{3\pi\log X}+E_{A,B}(X),$$

where

$$E_{A,B}(X) \ll \frac{X}{B\log X}+\frac{X}{A\log X}+\frac{X^{9/8}\log^3 X}{\sqrt{AB}}+\frac{(A+B)X^{3/4}\log^2 X}{AB}+\frac{X^{1/4}}{\log^2 X}.$$

The same result holds for counting trailing primes on average. Since extremal primes are the union of champion primes and trailing primes, we get an nearly identical result when counting extremal primes on average, where the only change is a constant of $16/3\pi$ in the asymptotic.

26 Oct 2017 17:44:55 PDT
Version 2 - Submitted to Acta Arith.

Corollary 1.5. *Taking $A, B > X^{3/4}\log X$ and the product $AB > X^{7/4}\log^{10} X$ in the previous theorem and letting $X \to \infty$ gives*

$$\frac{1}{4AB}\sum_{\substack{|a|\leq A\\ |b|\leq B}} \pi^{\text{Champ}}_{E(a,b)}(X) \sim \frac{8X^{1/4}}{3\pi \log X}.$$

The same result holds for counting trailing primes on average. For counting extremal primes on average the asymptotic is $\frac{16}{3\pi}\cdot X^{1/4}/\log X$.

It is worth pointing out that the contribution of complex multiplication curves to this average is negligible in comparison to the main term because CM curves have density 0 in the set of all elliptic curves (see [Jam04, Section 1] for a precise argument). In this way the significantly larger asymptotic in the CM case seen in Theorem 1.3 is not influential in Theorem 1.4, and therefore we may indeed interpret our result as the average over non-CM curves.

Using Theorem 1.4, we can prove a result on the variance of $\pi^{\text{Champ}}_{E(a,b)}(X)$ from the average $\frac{8}{3\pi}\cdot X^{1/4}/\log X$.

Theorem 1.6. *Upon taking $A, B > X\log^6 X$ and $AB > X^3\log^7 X$, we have*

$$\frac{1}{4AB}\sum_{\substack{|a|\leq A\\ |b|\leq B}} \left|\pi^{\text{Champ}}_{E(a,b)}(X) - \frac{8X^{1/4}}{3\pi\log X}\right|^2 \ll \frac{X^{1/2}}{\log^3 X}.$$

The Turán normal order method (see [CM05, Ch. 3]) subsequently yields the following corollary.

Corollary 1.7. *Fix $d > 1$. Upon taking A and B according to the hypotheses in Theorem 1.6, then with at most $\mathrm{O}(AB/\log^d X)$ exceptions, every model $E(a,b)$ with $|a| \leq A$ and $|b| \leq B$ satisfies the inequality*

$$\left|\pi^{\text{Champ}}_{E(a,b)}(X) - \frac{8X^{1/4}}{3\pi\log X}\right| < \frac{X^{1/4}}{\log^{3/2} X}.$$

Throughout the paper we will use the following notation. For coprime integers q and b, define

$$\begin{aligned} E(y,h;q,b) &:= \sum_{\substack{y<p\leq y+h\\ p\equiv b \pmod q}} \log p - \frac{h}{\phi(q)},\\ E(y,h;q) &:= \max_{(b,q)=1} |E(y,h;q,b)|. \end{aligned}$$

For a real Dirichlet character χ, we let

$$\begin{aligned} L(1,\chi) &:= \sum_{n\geq 1}\frac{\chi(n)}{n} = \prod_{\ell}\left(1-\frac{\chi(\ell)}{\ell}\right)^{-1}\\ L(1,\chi;w) &:= \prod_{\ell\leq w}\left(1-\frac{\chi(\ell)}{\ell}\right)^{-1}. \end{aligned}$$

26 Oct 2017 17:44:55 PDT
Version 2 - Submitted to Acta Arith.

We remark that using Mertens' third theorem, we obtain the bound

$$L(1,\chi;w) \ll \log w. \tag{1.1}$$

We also set $\Delta_p := [|2\sqrt{p}|]^2 - 4p$, $\Delta_{p,k} := k^2 - 4p$, and $P^+(n)$ to denote the largest prime factor of n.

The proof of Theorem 1.4 is similar to related problems in this area, such as in [FM96, DP99, Jam05, Par15, CDKS16]. In Section 2, we partition the box of elliptic curves by isomorphism class over $\mathbb{F}_p$ and appeal to a result of Deuring (see [Deu41]) which allows us to count the number of isomorphism classes of elliptic curves over $\mathbb{F}_p$ with precisely $p+1-r$ points for any integer r in terms of Hurwitz class numbers. As a result we reduce Theorem 1.4 to Theorem 2.1, which is a statement about primes in many arithmetic progressions. In Section 3, we give an outline of Theorem 2.1 conditional on technical results proved in Sections 4 and 5. In Section 6 we prove Theorem 1.6.

The main difficulty in this work arises from the need to count primes in short arithmetic progressions – arithmetic progressions so short that not even the generalized Riemann hypothesis would suffice. However since in this work we average over many short arithmetic progressions, we are able to invoke the following result from [Kou15].

Lemma 1.8 (Koukoulopoulos)**.** *Fix $\epsilon > 0$ and $R \geq 1$. For $2 \leq h \leq x$ and $1 \leq Q^2 \leq h/x^{1/6+\epsilon}$, we have*

$$\int_x^{2x} \sum_{q \leq Q} E(y,h;q)\, \mathrm{d}y \ll \frac{hx}{\log^R x}.$$

We also employ the following result from [CDKS16, Lemmas 2.2 and 2.3], which allows us to truncate $L(1,\chi)$ for most Dirichlet characters χ in a way conducive to applying Lemma 1.8.

Lemma 1.9. *Let $\alpha \geq 1$ and $H \geq 3$. For convenience in notation, set $z := \log H$. There is a set of integers $\mathcal{E}_\alpha(H) \subset [1,H]$ of size at most $H^{2/\alpha}$ such that if χ is a Dirichlet character of conductor $q \leq H$ not in $\mathcal{E}_\alpha(H)$, then*

$$L(1,\chi) = L(1,\chi; z^{8\alpha^2}) \left[1 + \mathrm{O}\left(\frac{1}{z^\alpha}\right)\right].$$

Moreover, for any $u \geq 1$ and $w \geq 10$, we have

$$L(1,\chi;w) = \sum_{\substack{n \leq w^u \\ P^+(n) \leq w}} \frac{\chi(n)}{n} + \mathrm{O}\left(\frac{\log w}{e^u}\right).$$

2. From Counting Curves to Counting Primes

We first write

$$\frac{1}{4AB} \sum_{\substack{|a| \leq A \\ |b| \leq B}} \pi^{\mathrm{Champ}}_{E(a,b)}(X) = \frac{1}{4AB} \sum_{3<p<X} N(A,B;p,-[|2\sqrt{p}|]), \tag{2.1}$$

where

$$N(A,B;p,r) := \#\left\{|a| \le A; |b| \le B : a_p(E(a,b)) = r\right\}.$$

For convenience, set $\Delta_p = \left[\left|2\sqrt{p}\right|\right]^2 - 4p$. Based on a classical result of Deuring, we cite a result of Baier (see [Bai07, Equation 4.5]) that gives

$$N(A,B;p,[|2\sqrt{p}|]) = \frac{2ABH(\Delta_p)}{p} + ERR_{A,B}(p),$$

where $ERR_{A,B}(p)$ is asymptotically bounded above by

(2.2)
$$\frac{AB}{p} + \frac{ABH(\Delta_p)}{p^2} + A + B + (ABH(\Delta_p))^{1/2}\log^3 p + \frac{(A+B)H(\Delta_p)}{p^{1/2}}\log p,$$

and where $H(d)$ is the Hurwitz class number associated to the discriminant d. We remark that from this result and the definition of Δ_p, we see that $N(A,B;p,\left[\left|2\sqrt{p}\right|\right]) = N(A,B;p,-\left[\left|2\sqrt{p}\right|\right])$. In this way, studying champion primes and trailing primes will each lead to an identical asymptotic.

Using the class number formula and the explicit formula for the Hurwitz class number we have the useful expression

$$H(\Delta_p) = \frac{1}{\pi}\sum_{\substack{f^2|\Delta_p \\ \frac{\Delta_p}{f^2}\equiv 0,1 \pmod 4}} \frac{\sqrt{|\Delta_p|}}{f} L\left(1, \chi_{\frac{\Delta_p}{f^2}}\right), \tag{2.3}$$

where $\chi_d := (d|\cdot)$ is the quadratic character associated with the discriminant d. We note the convexity bound

$$L(1,\chi_{\Delta_p/f^2}) \ll \log p \tag{2.4}$$

from [Lou93] and also convey the bounds

$$|\Delta_p| \le 4\sqrt{p}, \tag{2.5}$$

$$H(\Delta_p) \ll p^{1/4}\log^2 p. \tag{2.6}$$

With these estimates, the error term in (2.2) satisfies

$$ERR_{A,B}(p) \ll \frac{AB}{p} + \frac{AB\log^2 p}{p^{7/4}} + A + B + \sqrt{AB}p^{1/8}\log^4 p + \frac{(A+B)\log^3 p}{p^{1/4}}.$$

Substituting this work back into (2.1) gives

$$\frac{1}{4AB}\sum_{\substack{|a|\le A \\ |b|\le B}} \pi^{\text{Champ}}_{E(a,b)}(X) = \sum_{3<p<X}\frac{H(\Delta_p)}{2p} + E_{A,B}(X), \tag{2.7}$$

where $E_{A,B}(X)$ is asymptotically bounded above by

$$\log\log X + \frac{X}{B\log X} + \frac{X}{A\log X} + \frac{X^{9/8}\log^3 X}{\sqrt{AB}} + \frac{(A+B)X^{3/4}\log^2 X}{AB}.$$

We now focus solely on the main term of (2.7), which in conjunction with (2.3) is

$$\frac{1}{2\pi}\sum_{3<p<X}\frac{\sqrt{|\Delta_p|}}{p}\sum_{\substack{f^2\mid\Delta_p\\ \frac{\Delta_p}{f^2}\equiv 0,1 \pmod 4}}\frac{1}{f}L\left(1,\chi_{\frac{\Delta_p}{f^2}}\right). \tag{2.8}$$

As a function of a positive real variable t, the function $|\Delta_t| = 4t - [|2\sqrt{t}|]^2$ is sawtooth; it has zeros whenever t or $t/4$ is a square and is linear with slope 4 between these zeroes. With this in mind we define intervals

$$I_k := \left[\frac{k^2}{4}, \frac{(k+1)^2}{4}\right),$$

where we note that for $t \in I_k$ we have $|\Delta_t| = |\Delta_{t,k}| := |k^2 - 4t|$. Furthermore, since we are only concerned with the primes in the real interval $(3, X)$, it suffices to look at the union of intervals I_k from $k = 4$ to $k = [|2\sqrt{X}|]$. Partitioning in this manner allows us to write the quantity from (2.8) as

$$\frac{1}{2\pi}\sum_{3<k<2X^{1/2}}\sum_{p\in I_k}\frac{\sqrt{|\Delta_{p,k}|}}{p}\sum_{\substack{f^2\mid\Delta_{p,k}\\ \frac{\Delta_{p,k}}{f^2}\equiv 0,1 \pmod 4}}\frac{1}{f}L\left(1,\chi_{\frac{\Delta_{p,k}}{f^2}}\right) + \mathrm{O}(X^{1/8}\log^2 X), \tag{2.9}$$

where the error term arises from potential over counting in the interval $I_{[|2\sqrt{X}|]}$ since the parameter $2\sqrt{X}$ may not be an integer. It can be estimated naively using (2.4) and (2.5).

Switching the order of summations in (2.9) yields

$$\frac{1}{2\pi}\sum_{3<k<2X^{1/2}}\sum_{f\le\sqrt{2k+2}}\frac{1}{f}\sum_{p\in\mathcal{S}_f(I_k)}\frac{\sqrt{|\Delta_{p,k}|}}{p}L\left(1,\chi_{\frac{\Delta_{p,k}}{f^2}}\right) + \mathrm{O}(X^{1/8}\log^2 X), \tag{2.10}$$

where we have defined

$$\mathcal{S}_f(I_k) := \left\{p \in I_k : \Delta_{p,k} \equiv 0 \pmod{f^2}; \frac{\Delta_{p,k}}{f^2} \equiv 0, 1 \pmod 4\right\}.$$

The upper bound of $f \le \sqrt{2k+2}$ arises because if $f > \sqrt{2k+2}$ then the set $\mathcal{S}_f(I_k)$ is empty. More explicitly, a prime $p \in \mathcal{S}_f(I_k)$ only if $f^2 \mid \Delta_{p,k}$. As $\Delta_{p,k} \le 4\sqrt{p} \le 2(k+1)$, we see that $f^2 \mid \Delta_{p,k}$ occurs only if $f \le \sqrt{2k+2}$.

Lastly, we split the sum over $3 < k < 2X^{1/2}$ dyadically. The proof of the next statement spans the remainder of the paper.

Theorem 2.1. *For any real number $U \ge 4$, set*

$$\mathfrak{D}(U) := \sum_{U\le k<2U}\sum_{f\le\sqrt{2k+2}}\frac{1}{f}\sum_{p\in\mathcal{S}_f(I_k)}\frac{\sqrt{|\Delta_{p,k}|}}{p}L\left(1,\chi_{\frac{\Delta_{p,k}}{f^2}}\right).$$

We have that

$$\mathfrak{D}(U) = \frac{2\sqrt{2}}{3}\int_U^{2U}\frac{\mathrm{d}t}{t^{1/2}\log t} + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

Putting the result of Theorem 2.1 into (2.10) and integrating gives

$$\frac{\sqrt{2}}{3\pi}\int_4^{2X^{1/2}} \frac{\mathrm{d}t}{t^{1/2}\log t} + \mathrm{O}\left(\frac{X^{1/4}}{\log^2 X}\right) = \frac{8X^{1/4}}{3\pi\log X} + \mathrm{O}\left(\frac{X^{1/4}}{\log^2 X}\right),$$

which in the context of (2.7) proves Theorem 1.4 conditional on Theorem 2.1.

3. A Proof of Theorem 2.1

In what follows we let χ denote the quadratic character $\chi_{\Delta_{p,k}/f^2} = (\Delta p, k/f^2|\cdot)$. We begin by noting that for $p \in I_k$, a Taylor series approximation gives

$$\frac{1}{p\log p} = \frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}} + \mathrm{O}\left(\frac{1}{k^3\log k}\right).$$

Using this estimate gives

(3.1)
$$\mathfrak{D}(U) = \sum_{U\le k<2U} \frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}} \sum_{f\le\sqrt{2k+2}} \frac{1}{f} \sum_{p\in\mathcal{S}_f(I_k)} \sqrt{|\Delta_{p,k}|}\log p\cdot L\left(1,\chi\right) + \mathrm{O}(\log^2 U),$$

where the error term was bounded using (2.4) and (2.5).

We truncate the sum over integers $f \le \sqrt{2k+2} \le 3\sqrt{U}$ at a parameter $F := F(U)$ to be determined later. Using (2.4), (2.5), and the naive bound $\#\mathcal{S}_f(I_k) \ll k/f^2$ to bound the tail, the inner sums over $f \le \sqrt{2k+1}$ and $p \in \mathcal{S}_f(I_k)$ of (3.1) contribute

$$\sum_{f\le F}\frac{1}{f}\sum_{p\in\mathcal{S}_f(I_k)}\sqrt{|\Delta_{p,k}|}\log p\cdot L(1,\chi) + \mathrm{O}\left(\frac{k^{3/2}\log^2 k}{F^2}\right).$$

Therefore upon taking $F := \log^{3/2} U$ we have

(3.2)
$$\mathfrak{D}(U) = \sum_{U\le k<2U} \frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}} \sum_{f\le F} \frac{1}{f} \sum_{p\in\mathcal{S}_f(I_k)} \sqrt{|\Delta_{p,k}|}\log p\cdot L\left(1,\chi\right) + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

Of interest is to replace the special L-value $L(1,\chi)$ with an appropriate truncated L-value $L(1,\chi;w)$ for most characters χ. From the theory of quadratic characters we know that the conductor of $\chi := (\Delta_{p,k}/f^2|\cdot)$ is either $|\Delta_{p,k}/f^2|$ or $4|\Delta_{p,k}/f^2|$. Therefore as $U < k < 2U$, we may use (2.5) to bound the conductor of χ, denoted N_χ, above by $17U$.

Set $H := 17U$, let $\alpha = 4$, and let $\mathcal{E}_4(H)$ be the set of exceptional integers guaranteed by Lemma 1.9. For convenience, set $z := \log H$. For a prime $p \in \mathcal{S}_f(I_k)$, if $N_\chi \in \mathcal{E}_4(H)$ then by (2.4), we have

$$L(1,\chi) - L(1,\chi;z^{128}) \ll \log p \ll \log U.$$

On the other hand, if $N_\chi \notin \mathcal{E}_4(H)$ then by Lemma 1.9 and (1.1) we have

$$L(1,\chi) - L(1,\chi,z^{128}) \ll \frac{L(1,\chi;z^{128})}{z^4} \ll \frac{\log z}{z^4} \ll \frac{1}{\log^3 U}.$$

Lastly we note that as p runs through $\mathcal{S}_f(I_k)$ and we compute the conductor of χ, we never encounter the same conductor more than twice. To see this, we explicitly bound

$$\sum_{\substack{p\in\mathcal{S}_f(I_k)\\ N_\chi=e}} 1 \le \sum_{\substack{m\in I_k\\ (4m-k^2)/f^2=e}} 1 + \sum_{\substack{m\in I_k\\ 4(4m-k^2)/f^2=e}} 1 \le 2.$$

Therefore, for a fixed integer k satisfying $U \le k < 2U$ and a fixed integer $f < F$ we have

$$\begin{aligned}\sum_{p\in\mathcal{S}_f(I_k)} &\sqrt{|\Delta_{p,k}|}\log p\cdot\left[L\left(1,\chi\right) - L\left(1,\chi;z^{128}\right)\right]\\ &\ll \sqrt{U}\log U\left[\sum_{\substack{p\in\mathcal{S}_f(I_k)\\ N_\chi\in\mathcal{E}_4(H)}}\log U + \sum_{\substack{p\in\mathcal{S}_f(I_k)\\ N_\chi\notin\mathcal{E}_4(H)}}\frac{1}{\log^3 U}\right]\\ &\ll \sqrt{U}\log U\left[H^{1/2}\log U + \frac{U}{\log^3 U}\right].\end{aligned}$$

Recalling that $H \ll U$ gives the entire error above is $\mathrm{O}(U^{3/2}/\log^2 U)$. Therefore we have shown $\mathfrak{D}(U)$ to be equal to

(3.3)
$$\sum_{U\le k<2U}\frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}}\sum_{f\le F}\frac{1}{f}\sum_{p\in\mathcal{S}_f(I_k)}\sqrt{|\Delta_{p,k}|}\log p\cdot L\left(1,\chi;z^{128}\right) + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

With $z = \log(17U)$, one can check that $z^{128} \ge 10$ for any $U \ge 1$. Therefore by Lemma 1.9, with $v := 4\log\log U$, we have

$$L\left(1,\chi;z^{128}\right) = \sum_{\substack{n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{\chi(n)}{n} + \mathrm{O}\left(\frac{1}{\log^3 U}\right).$$

As a result we see that $\mathfrak{D}(U)$ is equal to

(3.4)
$$\sum_{U\le k<2U}\frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}}\sum_{\substack{f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{1}{nf}\sum_{p\in\mathcal{S}_f(I_k)}\sqrt{|\Delta_{p,k}|}\log p\cdot\chi(n) + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

For now we solely investigate the main term of (3.4). As discussed previously, $\chi(n)$ is the Kronecker symbol $\left(\frac{\Delta_{p,k}/f^2}{n}\right)$. Using its $4n$-periodicity, we can write

$$\chi(n) = \left(\frac{a}{n}\right) \quad \text{for } a \equiv \Delta_{p,k}/f^2 \pmod{4n}.$$

Furthermore, the conditions $p \in \mathcal{S}_f(I_k)$ and $\Delta_{p,k}/f^2 \equiv a \pmod{4n}$ are equivalent to $p \in I_k$, $4p \equiv (k^2 - af^2) \pmod{4nf^2}$, and $a \equiv 0, 1 \pmod 4$. In

this way, the main term of (3.4) is

$$\sum_{U\le k<2U}\frac{1}{\frac{k^2}{4}\log\frac{k^2}{4}}\sum_{\substack{f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{1}{nf}\sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z}\\ a\equiv 0,1 \pmod 4\\ (k^2-af^2,4nf^2)=4}}\left(\frac{a}{n}\right)\sum_{\substack{p\in I_k\\ p\equiv\frac{k^2-af^2}{4} \pmod{nf^2}}}\sqrt{|\Delta_{p,k}|}\log p, \tag{3.5}$$

where the condition $(k^2-af^2, 4nf^2)=4$ is a necessary condition when the inner sum over primes is non-zero.

The innermost sum counts primes in a short arithmetic progressions with an awkward weighting function. Using the method shown in [CDKS16, Lemma 7.1], the following result is proved in Section 4.

Lemma 3.1. *Let b, q be coprime integers, and fix an integer k satisfying $U\le k<2U$. Define*

$$\Lambda(k;q,b):=\sum_{\substack{p\in I_k\\ p\equiv b \pmod q}}\sqrt{|\Delta_{p,k}|}\log p.$$

For any $q\le h\le U/4$, we have

$$\Lambda(k;q,b)=\frac{(2k+1)^{3/2}}{6\phi(q)}+\mathrm{O}\left(\frac{U^{1/2}}{h}\int_{I_k}|E(y,h;q,b)|\ \mathrm{d}y+\frac{hU^{1/2}\log U}{q}\right).$$

Applying the result of Lemma 3.1 to (3.4) and (3.5) while also rearranging some finite sums gives the expression

$$\begin{aligned}\mathfrak{D}(U) &= \frac{1}{6}\sum_{U\le k<2U}\frac{(2k+1)^{3/2}}{\frac{k^2}{4}\log\frac{k^2}{4}}\left[\sum_{\substack{f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{1}{nf\phi(nf^2)}\sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z}\\ a\equiv 0,1 \pmod 4\\ (k^2-af^2,4nf^2)=4}}\left(\frac{a}{n}\right)\right]\\ &\quad+\mathrm{O}\left(E_1(U,h)+E_2(U,h)+\frac{U^{1/2}}{\log^2 U}\right),\end{aligned} \tag{3.6}$$

where

$$E_1(U,h)\ll\frac{U^{1/2}}{h}\sum_{\substack{U\le k<2U\\ f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{1}{k^2nf\log k}\sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z}\\ a\equiv 0,1 \pmod 4\\ (k^2-af^2,4nf^2)=4}}\int_{I_k}\left|E\left(y,h;nf^2,\frac{k^2-af^2}{4}\right)\right|\mathrm{d}y,$$

$$E_2(U,h)\ll h\sqrt{U}\log U\sum_{U\le k<2U}\frac{1}{k^2\log k}\sum_{\substack{f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}}\frac{1}{n^2f^3}\sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z}\\ a\equiv 0,1 \pmod 4\\ (k^2-af^2,4nf^2)=4}}1.$$

Since $v:=4\log\log U$ and $z:=\log(17U)$, $E_2(U,h)$ can be asymptotically bounded above by

$$\frac{h}{\sqrt{U}}\sum_{n\le z^{128v}}\frac{1}{n}\sum_{f<F}\frac{1}{f^3}\ \ll\ \frac{hv\log z}{\sqrt{U}}\ \ll\ \frac{h(\log\log U)^2}{\sqrt{U}},$$

so if we take $h \ll U/\log^4 U$ then this error is $\mathrm{O}(U^{1/2}/\log^2 U)$.

The $E_1(U,h)$ term can be bounded using Lemma 1.8. As we have used before, for $U \le k < 2U$ the function $1/(k^2 \log k) = \mathrm{O}(1/(U^2 \log U))$. Therefore $E_1(U,h)$ is asymptotically bounded above by

$$\frac{1}{hU^{3/2}\log U} \sum_{\substack{U\le k<2U \\ f<F \\ n\le z^{128v} \\ P^+(n)\le z^{128}}} \frac{1}{nf} \sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z} \\ a\equiv 0,1 \pmod 4 \\ (k^2-af^2,4nf^2)=4}} \int_{I_k} \left| E\left(y,h;nf^2,\frac{k^2-af^2}{4}\right)\right| \,\mathrm{d}y.$$

Since $|E(y,h;q,b)| \le \max_{(b,q)=1} |E(y,h;q,b)| =: E(y,h;q)$, we obtain

$$E_1(U,h) \ll \frac{1}{hU^{3/2}\log U} \sum_{U\le k<2U} \sum_{\substack{f<F \\ n\le z^{128v} \\ P^+(n)\le z^{128}}} \frac{1}{f} \int_{I_k} E\left(y,h;nf^2\right) \,\mathrm{d}y,$$

where now as all summands are positive we may relax the conditions on the sum over integers n to give

$$E_1(U,h) \ll \frac{1}{hU^{3/2}\log U} \sum_{U\le k<2U} \sum_{\substack{f<F \\ n\le z^{128v}}} \frac{1}{f} \int_{I_k} E\left(y,h;nf^2\right) . \,\mathrm{d}y$$

Switching sums and integrals is valid as these are all finite quantities. Furthermore, we recall that the intervals $I_U, I_{U+1}, \dots, I_{2U}$ partition the real interval $[U^2/4, U^2]$. Lastly, upon fixing an f as n runs through the interval $n \le z^{128v}$ we see moduli of the form $nf^2 \le z^{128v} f^2$. Since all summands are positive we may extend the sum to all moduli $q \le z^{128v} f^2$. With all these observations, we continue with

$$E_1(U,h) \ll \frac{1}{hU^{3/2}\log U} \sum_{f<F} \frac{1}{f} \sum_{q<z^{128v}f^2} \int_{U^2/4}^{U^2} E\left(y,h;q\right) \,\mathrm{d}y.$$

We recall that $h \ll U/\log^4 U$, $F = \log^{3/2} U$, $z = \log(17U)$, and $v = 4\log\log U$. For any $f < F$, $(z^{128v} f^2)^2 \ll U^\delta$ for any $\delta > 0$, and therefore the condition $(z^{128v}F^2)^2 \le h/U^{1/6+\epsilon}$ for some $\epsilon > 0$ in Lemma 1.8 holds comfortably by choosing $h = CU/\log^4 U$ with a suitably large constant C. Therefore applying Lemma 1.8 with $R = 2$ to this quantity gives

$$E_1(U,h) \ll \frac{\log F}{hU^{3/2}\log U} \cdot \frac{hU^2}{\log^2 U} \ll \frac{U^{1/2}}{\log^2 U}.$$

With both $E_1(U,h)$ and $E_2(U,h)$ estimated, from (3.6) we have shown that $\mathfrak{D}(U)$ is

(3.7)

$$\frac{1}{6} \sum_{U\le k<2U} \frac{(2k+1)^{3/2}}{\frac{k^2}{4}\log\frac{k^2}{4}} \left[\sum_{\substack{f\le F \\ n\le z^{128v} \\ P^+(n)\le z^{128}}} \frac{1}{nf\phi(nf^2)} \sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z} \\ a\equiv 0,1 \pmod 4 \\ (k^2-af^2,4nf^2)=4}} \left(\frac{a}{n}\right) \right] + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

The quantity in square brackets in (3.7) is almost identical to something investigated in [DP99, Section 3]; the only difference in the two terms is our additional constraint of $P^+(n) \le z^{128}$. However, seeing as the analysis is essential identical we omit the proof and give full credit to the authors of that paper for the fact that

$$\sum_{\substack{f\le F\\ n\le z^{128v}\\ P^+(n)\le z^{128}}} \frac{1}{nf\phi(nf^2)} \sum_{\substack{a\in\mathbb{Z}/4n\mathbb{Z}\\ a\equiv 0,1 \pmod 4\\ (k^2-af^2,4nf^2)=4}} \left(\frac{a}{n}\right) = C\cdot C_k + \mathrm{O}\left(\frac{1}{F^2}+\frac{1}{z^{64v}}+\frac{1}{z^{64}}\right),$$

where

$$C\cdot C_k := \prod_{\ell} \frac{\ell(\ell^2-\ell-1)}{(\ell-1)(\ell^2-1)} \cdot \prod_{\ell\mid k} \frac{\ell(\ell-1)}{\ell^2-\ell-1}.$$

Note that our choice of parameters $F = \log^{3/2} U$, $v = 4\log\log U$, and $z = \log 17U$ gives that the entire error above is $\mathrm{O}(1/\log^3 U)$. As a result, we apply this result to (3.7) to obtain

$$\mathfrak{D}(U) = \frac{C}{6} \sum_{U\le k<2U} \frac{(2k+1)^{3/2}}{\frac{k^2}{4}\log\frac{k^2}{4}} \cdot C(k) + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right).$$

Since $U \le k < 2U$, Taylor series approximations allow us to write

$$\begin{aligned} (2k+1)^{3/2} &= 2\sqrt{2}k^{3/2} + \mathrm{O}(k^{1/2}),\\ \frac{1}{\log\frac{k^2}{4}} &= \frac{1}{2\log k} + \mathrm{O}\left(\frac{1}{\log^2 k}\right). \end{aligned}$$

Therefore upon estimating the resulting error terms we have

$$\mathfrak{D}(U) = \frac{2\sqrt{2}C}{3} \sum_{U\le k<2U} \frac{C(k)}{k^{1/2}\log k} + \mathrm{O}\left(\frac{U^{1/2}}{\log^2 U}\right). \tag{3.8}$$

The partial sums of $C(k)$ are studied in Section 5.

Lemma 3.2. *Let $U < K$ be positive real numbers. We have*

$$S(U,K) := \sum_{U\le k<K} C(k) = C^{-1}(K-U) + \mathrm{O}(\log K),$$

where

$$C^{-1} := \prod_{\ell} \left(1 + \frac{1}{\ell^3-\ell^2-\ell}\right).$$

Since $CC^{-1} = 1$, applying partial summation and then Lemma 3.2 to the main term of (3.8) gives

$$\begin{aligned} &\frac{2\sqrt{2}C}{3}\left[\frac{S(1,2U)}{(2U)^{1/2}\log 2U} - \frac{S(1,U)}{(U)^{1/2}\log U} - \int_U^{2U} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{t^{1/2}\log t}\right) S(1,t)\ \mathrm{d}t\right]\\ &= \frac{2\sqrt{2}}{3}\left[\frac{2U}{(2U)^{1/2}\log 2U} - \frac{U}{(U)^{1/2}\log U} - \int_U^{2U} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{t^{1/2}\log t}\right) t\ \mathrm{d}t\right] \end{aligned}$$

with a negligible error. Integrating by parts gives this quantity as

$$\frac{2\sqrt{2}}{3}\int_U^{2U}\frac{\mathrm{d}t}{t^{1/2}\log t},$$

which completes the proof of Theorem 2.1 conditional on the lemmas used in the proof.

4. A Proof of Lemma 3.1

In this section we study the quantity

$$\Lambda(k;q,b) := \sum_{\substack{p\in I_k\\ p\equiv b \pmod q}} \sqrt{|\Delta_{p,k}|}\cdot\log p,$$

where $U\le k<2U$ and the integers b and q are coprime. We also define the notation

$$\begin{aligned}\ell_k &:= \frac{2k+1}{4};\\ I_k^+ &:= \frac{(k+1)^2}{4};\\ I_k^- &:= \frac{k^2}{4}.\end{aligned}$$

We begin by peeling off a small amount from each end of the interval I_k. Let $h\le U/4$. Since $U\le k<2U$, we know that $I_k^-+2h\le I_k^-+U/2<I_k^+$. Using the naive bound $\#\{p\in(y,y+z):p\equiv b \pmod q\}\ll z/q$, we have

$$\Lambda(k;q,b)=\sum_{\substack{I_k^-+h<p\le I_k^+-h\\ p\equiv b \pmod q}}\sqrt{|\Delta_{p,k}|}\cdot\log p+E_1, \tag{4.1}$$

where

$$E_1\ll\frac{h\sqrt{k}\log k}{q}\ll\frac{h\sqrt{U}\log U}{q}:=F. \tag{4.2}$$

For any prime p satisfying $I_k^-+h<p\le I_k^+-h$, we can write $p=I_k^-+p_0\cdot\ell_k$ for some real number p_0 satisfying $\frac{h}{\ell_k}<p_0\le 1-\frac{h}{\ell_k}$. Therefore, for such primes we have

$$\sqrt{|\Delta_{p,k}|}=\sqrt{2k+1}\cdot\sqrt{p_0}.$$

Set $\eta:=h/\ell_k$. For any $t=t_0+\mathrm{O}(\eta)$, a Taylor series approximation gives

$$\int_{t_0-\eta}^{t_0}\sqrt{t}\ \mathrm{d}t=\eta\sqrt{t_0}+\mathrm{O}\left(\frac{\eta^2}{\sqrt{t_0}}\right).$$

Therefore for the primes described above, we have

$$\sqrt{|\Delta_{p,k}|}=\frac{(2k+1)^{3/2}}{4h}\int_{\frac{p-I_k^-}{\ell_k}-\frac{h}{\ell_k}}^{\frac{p-I_k^-}{\ell_k}}\sqrt{t}\ \mathrm{d}t+\mathrm{O}\left(\frac{\eta}{\sqrt{\frac{p-I_k^-}{\ell_k}}}\right).$$

Upon putting this work into (4.1), we obtain

$$\Lambda(k;q,b) = \frac{(2k+1)^{3/2}}{4h} \sum_{\substack{I_k^- + h < p \le I_k^+ - h \\ p \equiv b \pmod q}} \log p \left[\int_{\frac{p-I_k^-}{\ell_k}}^{\frac{p-I_k^-}{\ell_k}} \sqrt{t}\, \mathrm{d}t \right] + \mathrm{O}(F) + E_2,$$

where

$$E_2 \ll \frac{h \log k}{\sqrt{k}} \sum_{\substack{I_k^- + h < p \le I_k^+ - h \\ p \equiv b \pmod q}} \frac{1}{\sqrt{p - I_k^-}} \ll \frac{\sqrt{hU} \log U}{q} \ll F. \tag{4.3}$$

Switching the sum and integral gives

$$\Lambda(k;q,b) = \frac{(2k+1)^{3/2}}{4h} \int_0^{1-\eta} \sqrt{t} \left[\sum_{\substack{I_k^- + h < p \le I_k^+ - h \\ I_k^- + t\ell_k < p < I_k^- + t\ell_k + h \\ p \equiv b \pmod q}} \log p \right] \mathrm{d}t + \mathrm{O}(F). \tag{4.4}$$

We aim to extend the limits of the integration to the full interval $[0,1]$. For t satisfying $\eta < t < 1 - 2\eta$, the first condition in the summation is implied by the second condition. For $t \in [0,1] \backslash (\eta, 1-2\eta)$, we have

$$\sqrt{t} \cdot \sum_{\substack{I_k^- + h < p \le I_k^+ - h \\ I_k^- + t\ell_k < p < I_k^- + t\ell_k + h \\ p \equiv b \pmod q}} \log p \le \sum_{\substack{I_k^- + t\ell_k < p < I_k^- + t\ell_k + h \\ p \equiv b \pmod q}} \log p \ll \frac{h \log k}{q},$$

and so the full contribution for $t \in [0,1] \backslash (\eta, 1-2\eta)$ is

$$\frac{(2k+1)^{3/2}}{4h} \int_{[0,1]\backslash(\eta,1-2\eta)} \frac{h \log k}{q}\, \mathrm{d}t \ll \frac{k^{1/2} h \log k}{q} \ll \frac{U^{1/2} h \log U}{q} \ll F.$$

Returning to (4.4), we now have the expression

$$\Lambda(k;q,b) = \frac{(2k+1)^{3/2}}{4h} \int_0^{1} \sqrt{t} \left[\sum_{\substack{I_k^- + t\ell_k < p < I_k^- + t\ell_k + h \\ p \equiv b \pmod q}} \log p \right] \mathrm{d}t + \mathrm{O}(F). \tag{4.5}$$

The sum above counts log-weighted primes in a short arithmetic progression. For an interval $(y, y+z]$ and a pair of coprime integers b and q, we define a quantity $E(y,z;q,b)$ such that

$$\sum_{\substack{y < p \le y+z \\ p \equiv b \pmod q}} \log p = \frac{z}{\phi(q)} + E(y,z;q,b).$$

Applying this substitution to the weighted sum of primes, we have

$$\Lambda(k;q,b) = \frac{(2k+1)^{3/2}}{4h} \int_0^{1} \sqrt{t} \left[\frac{h}{\phi(q)} + E(I_k^- + t\ell_k, h; q, b) \right] \mathrm{d}t + \mathrm{O}(F),$$

which is

$$\frac{(2k+1)^{3/2}}{6\phi(q)} + \frac{(2k+1)^{3/2}}{4h}\int_0^1 \sqrt{t}\cdot E(I_k^- + t\ell_k, h; q, b)\ \mathrm{d}t + \mathrm{O}(F).$$

Upon applying the change of variables $y := I_k^- + t\ell_k$, the second term is no larger in absolute value than

$$\frac{k^{3/2}}{h}\int_0^1 \left|E(I_k^- + t\ell_k, h; q, b)\right|\ \mathrm{d}t \ll \frac{U^{1/2}}{h}\int_{I_k} |E(y, h; q, b)|\ \mathrm{d}y,$$

which completes the proof.

5. A Proof of Lemma 3.2

In this section we study the quantity

$$S(U,K) := \sum_{U\le k<K} C(k) = \sum_{U\le k<K}\left(\prod_{\ell\mid k}\frac{\ell(\ell-1)}{\ell^2-\ell-1}\right),$$

however it suffices to prove

$$S(K) := \sum_{k<K} C(k) = C^{-1}K + \mathrm{O}(\log K), \tag{5.1}$$

where C^{-1} is defined in the statement of the lemma, as then $S(U,K) = S(K) - S(U)$ gives the desired result.

We note begin by noting that $C(k)$ is multiplicative, and so the identity

$$C(k) = \sum_{d\mid k}\mu^2(d)\prod_{\ell\mid d}\frac{1}{\ell^2-\ell-1}$$

can be checked by computing on prime powers. With this in tow, we sum over all $k < K$ to obtain

$$\begin{aligned}
S(K) &= \sum_{k<K} C(k)\\
&= \sum_{k<K}\sum_{d\mid k}\mu^2(d)\prod_{\ell\mid d}\frac{1}{\ell^2-\ell-1}\\
&= \sum_{de<K}\mu^2(d)\prod_{\ell\mid d}\frac{1}{\ell^2-\ell-1}\\
&= \sum_{d<K}\mu^2(d)\prod_{\ell\mid d}\frac{1}{\ell^2-\ell-1}\left[\sum_{e<K/d} 1\right]\\
&= K\sum_{d<K}\frac{\mu^2(d)}{d^3}\prod_{\ell\mid d}\frac{\ell^2}{\ell^2-\ell-1} - \sum_{d<K}\left\{\frac{K}{d}\right\}\frac{\mu^2(d)}{d^2}\prod_{\ell\mid d}\frac{\ell^2}{\ell^2-\ell-1}.
\end{aligned}$$

Let $\nu(d)$ count the number of prime divisors of d without multiplicity. We remark that since $1 < \ell^2/(\ell^2-\ell-1) \le 9/5$ for any prime $\ell > 2$, the

fractional part $\{K/d\} \leq 1$, and the arithmetic function $\nu(d) \leq \log d$, the second term above can be bounded above by

$$\sum_{d<K} \frac{\mu^2(d)}{d^2} \cdot \left[4 \cdot \left(\frac{9}{5}\right)^{\nu(d)} \right] \ll \sum_{d<K} \frac{\mu^2(d)}{d^2} \cdot \left(\frac{9}{5}\right)^{\log d} \ll \sum_{d<K} \frac{1}{d^2} \cdot d \ll \log K.$$

For the main term, we extend the sum and estimate the tail similarly as

$$K \sum_{d\geq K} \frac{\mu^2(d)}{d^3} \prod_{\ell|d} \frac{\ell^2}{\ell^2-\ell-1} \ll K \sum_{d\geq K} \frac{\mu^2(d)}{d^3} \cdot \left(\frac{9}{5}\right)^{\log d} \ll K \sum_{d\geq K} \frac{1}{d^2} \ll 1.$$

As a result, in full we have

$$\begin{aligned} S(K) &= K \sum_{d=1}^{\infty} \frac{\mu^2(d)}{d^3} \prod_{\ell|d} \frac{\ell^2}{\ell^2-\ell-1} + \mathrm{O}(\log K + 1) \\ &= K \prod_{l} \sum_{\alpha=0}^{\infty} \frac{\mu^2(\ell^\alpha)}{\ell^{3\alpha}} \prod_{p|\ell^\alpha} \frac{p^2}{p^2-p-1} + \mathrm{O}(\log K) \\ &= K \prod_{\ell} \left(1 + \frac{1}{\ell^3} \cdot \frac{\ell^2}{\ell^2-\ell-1}\right) + \mathrm{O}(\log K) \\ &= C^{-1} K + \mathrm{O}(\log K), \end{aligned}$$

which completes the proof of (5.1).

6. A Proof of Theorem 1.6

In this section we study the quantity

$$\mathcal{V} := \frac{1}{4AB} \sum_{\substack{|a|\leq A \\ |b|\leq B}} \left| \pi^{\mathrm{Champ}}_{E(a,b)}(X) - \frac{8X^{1/4}}{3\pi \log X} \right|^2.$$

Upon taking A and B according to the conditions in Corollary 1.5, we see that

$$\mu := \frac{1}{4AB} \sum_{\substack{|a|\leq A \\ |b|\leq B}} \pi^{\mathrm{Champ}}_{E(a,b)}(X) = \frac{8X^{1/4}}{3\pi \log X} + \mathrm{O}\left(\frac{X^{1/4}}{\log^2 X}\right).$$

Therefore applying the triangle-inequality to the definition of $\mathcal{V}$ we see that

$$\mathcal{V} \ll \frac{1}{4AB} \sum_{\substack{|a|\leq A \\ |b|\leq B}} \left| \pi^{\mathrm{Champ}}_{E(a,b)}(X) - \mu \right|^2 + \mathrm{O}\left(\frac{X^{1/2}}{\log^4 X}\right),$$

whereupon expanding the product we obtain

$$\mathcal{V} \ll \frac{1}{4AB} \sum_{\substack{|a|\leq A \\ |b|\leq B}} \left[\pi^{\mathrm{Champ}}_{E(a,b)}(X) \right]^2 - \mu^2 + \mathrm{O}\left(\frac{X^{1/2}}{\log^4 X}\right). \tag{6.1}$$

For now we focus solely on the first term of (6.1). We begin by writing

$$\left[\pi^{\text{Champ}}_{E(a,b)}(X)\right]^2 = \pi^{\text{Champ}}_{E(a,b)}(X) \quad + \sum_{\substack{3<p,q<X\\ p\neq q\\ a_p(E(a,b))=\left[\left|-2\sqrt{p}\right|\right]\\ a_q(E(a,b))=\left[\left|-2\sqrt{q}\right|\right]}} 1 \quad + \quad \mathrm{O}(1),$$

where the error term corresponds to the finite number of primes of bad reduction of $E(a,b)$. Upon averaging over $|a|\leq A$ and $|b|\leq B$ and switching the resulting finite sums this first term of (6.1) is

$$\mu \quad + \quad \frac{1}{4AB}\sum_{\substack{3<p,q<X\\ p\neq q}} \sum_{\substack{|a|\leq A\\ |b|\leq B\\ a_p(E(a,b))=\left[\left|-2\sqrt{p}\right|\right]\\ a_q(E(a,b))=\left[\left|-2\sqrt{q}\right|\right]}} 1 \quad + \quad \mathrm{O}(1).$$

For each isomorphism class $\tilde{E}_1/\mathbb{F}_p$ and $\tilde{E}_2/\mathbb{F}_q$, choose representative models $E_1(a,b)$ and $E_2(a,b)$. Partitioning the models by isomorphism class allows us to rewrite the above as

$$\mu \quad + \quad \frac{1}{4AB}\sum_{\substack{3<p,q<X\\ p\neq q}} \sum_{\substack{\tilde{E}_1/\mathbb{F}_p\\ a_p(E_1)=\left[\left|-2\sqrt{p}\right|\right]}} \sum_{\substack{\tilde{E}_2/\mathbb{F}_q\\ a_p(E_2)=\left[\left|-2\sqrt{q}\right|\right]}} \sum_{\substack{|a|\leq A\\ |b|\leq B\\ E(a,b)\cong_p E_1(a,b)\\ E(a,b)\cong_q E_2(a,b)}} 1 \quad + \quad \mathrm{O}(1).$$

The inner-most sum above counts models of elliptic curves that are $\mathbb{F}_p$-isomorphic to $E_1(a,b)$ and $\mathbb{F}_q$-isomorphic to $E_2(a,b)$. Specializing Lemma 2 of [JS11] to the field $K=\mathbb{Q}$ gives this inner sum as

$$\frac{4AB(p-1)(q-1)}{(pq)^2\#\mathrm{Aut}_p(E_1)\#\mathrm{Aut}_q(E_2)} + \mathrm{O}\left(\frac{4AB}{pq\min\{A,B\}} + (pq)^{1/2}\log^2 pq + \frac{A+B}{(pq)^{1/2}}\cdot\log pq\right).$$

As a result the main term of (6.1) is

$$\mu \quad + \quad \sum_{\substack{3<p,q<X\\ p\neq q}} \frac{(p-1)(q-1)}{(pq)^2}\cdot N(p)N(q) + \mathrm{O}\left(\frac{1}{\min\{A,B\}}\sum_{\substack{3<p,q<X\\ p\neq q}}\frac{1}{pq}\cdot N(p)N(q)\right) + \mathrm{O}\left(\frac{1}{4AB}\sum_{\substack{3<p,q<X\\ p\neq q}}(pq)^{1/2}\log pq\left[\log pq + \frac{A+B}{pq}\right]\cdot N(p)N(q)\right),$$

where $N(q)$ denotes the number of isomorphism classes of elliptic curves $\tilde{E}$ over $\mathbb{F}_p$ with trace $a_p(\tilde{E}) = [|-2\sqrt{p}|]$, weighted by $1/\#\mathrm{Aut}_p(\tilde{E})$. Using

Deuring's theorem from [Deu41], for $r^2 < 4p$ the number of $\mathbb{F}_p$-isomorphism classes of elliptic curves with precisely $p+1-r$ points is $H(r^2-4p)+\mathrm{O}(1)$. Since $\#\mathrm{Aut}_p(E_1) = 2$ for all but $10 = \mathrm{O}(1)$ isomorphism classes, the above simplifies to

$$\begin{aligned}\mu \;+\; & \frac{1}{4}\sum_{\substack{3<p,q<X\\ p\neq q}} \frac{(p-1)(q-1)}{(pq)^2} H(\Delta_p)H(\Delta_q) \\ & +\mathrm{O}\left(\frac{1}{\min\{A,B\}}\sum_{\substack{3<p,q<X\\ p\neq q}} \frac{H(\Delta_p)H(\Delta_q)}{pq}\right) \\ & +\mathrm{O}\left(\frac{1}{4AB}\sum_{\substack{3<p,q<X\\ p\neq q}} (pq)^{1/2}\log pq\left[\log pq + \frac{A+B}{pq}\right] H(\Delta_p)H(\Delta_q)\right),\end{aligned}$$

where we recall the notation $\Delta_t := \left[\left|2\sqrt{t}\right|\right]^2 - 4t$. Note from Theorem 1.4 that $\mu \ll X^{1/4}/\log X$. Using the bound for the Hurwitz class number in (2.6) implies the total contribution from the first error term above is asymptotically smaller than

$$\frac{1}{\min\{A,B\}}\sum_{\substack{3<p,q<X\\ p\neq q}} \frac{\log^2 p \log^2 q}{(pq)^{3/4}} \ll \frac{X^{1/2}\log^2 X}{\min\{A,B\}}, \tag{6.2}$$

while the second error term is less than

$$\begin{aligned}&\frac{1}{4AB}\sum_{\substack{3<p,q<X\\ p\neq q}} (pq)^{3/4}\log^5 pq\left[\log pq + \frac{A+B}{pq}\right] \\ &\qquad \ll \frac{X^{7/2}\log^4 X}{4AB} + X^{3/2}\log^3 X\left[\frac{1}{A}+\frac{1}{B}\right].\end{aligned} \tag{6.3}$$

Since $A, B > X\log^6 X$ and $AB > X^3\log^7 X$ the quantities in (6.2) and (6.3) are both bounded above by a constant multiple of $X^{1/2}/\log^3 X$. As a result the main term of (6.1) is

$$\left[\sum_{3<p<X}\frac{H(\Delta_p)}{2p}\right]^2 + \mathrm{O}\left(\frac{X^{1/2}}{\log^3 X}\right).$$

From (2.7) and Theorem 1.4, we may write

$$\left[\sum_{3<p<X}\frac{H(\Delta_p)}{2p}\right]^2 = \left[\mu + \mathrm{O}\left(\frac{X^{1/4}}{\log^2 X}\right)\right]^2 = \mu^2 + \mathrm{O}\left(\frac{X^{1/2}}{\log^3 X}\right).$$

Putting the previous two lines into (6.1) gives the desired result.

7. Acknowledgments

The authors would like to thank Ethan Smith for pointing out the result found in Lemma 1.8 as it was crucial to the completion of this research. This work was supported by an NSF Research Training Group (RTG) grant (DMS #1547399) promoting Coding Theory, Cryptography, and Number Theory at Clemson University.

References

[Bai07] Stephan Baier. The Lang-Trotter conjecture on average. *Journal of the Ramanujan Mathematical Society*, 22:299–314, 2007.

[BLGHT11] Tom Barnet-Lamb, David Geraghty, Michael Harris, and Richard Taylor. A family of Calabi–Yau varieties and potential automorphy II. *Publications of the Research Institute for Mathematical Sciences*, 47(1):29–98, 2011.

[CDKS16] Vorrapan Chandee, Chantal David, Dimitris Koukoulopoulos, and Ethan Smith. Frequency of elliptic curve groups over prime finite fields. *Canadian Journal of Mathematics*, 68(4):721–761, 2016.

[CHT08] Laurent Clozel, Michael Harris, and Richard Taylor. Automorphy for some ℓ-adic lifts of automorphic mod ℓ Galois representations. *Publications mathématiques*, 108(1):1, 2008.

[CM05] Alina Cojocaru and M. Ram Murty. *An Introduction to Sieve Methods and Their Applications*. Cambridge University Press, 2005.

[Deu41] Max Deuring. Die typen der multiplikatorenringe elliptischer funktionenkörper. In *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, volume 14, pages 197–272, 1941.

[DP99] Chantal David and Francesco Pappalardi. Average Frobenius distributions of elliptic curves. *International Math Research Notices*, (4):165 – 183, 1999.

[FM96] Etienne Fouvry and M Ram Murty. On the distribution of supersingular primes. *Canadian Journal of Mathematics*, 48(1):81–104, 1996.

[HJX14] Jason Hedetniemi, Kevin James, and Hui Xue. Champion primes for elliptic curves over fields of prime order. *INTEGERS*, 14, 2014.

[HSBT10] Michael Harris, Nick Shepherd-Barron, and Richard Taylor. A family of Calabi-Yau varieties and potential automorphy. *Annals of Mathematics*, pages 779–813, 2010.

[Jam04] Kevin James. Average Frobenius distributions for elliptic curves with 3-torsion. *Journal of Number Theory*, 109(2):278–298, 2004.

[Jam05] Kevin James. Averaging special values of Dirichlet L-series. *The Ramanujan Journal*, 10(1):75–87, 2005.

[JP17] Kevin James and Paul Pollack. Extremal primes for elliptic curves with complex multiplication. *Journal of Number Theory*, 172:383–391, 2017.

[JS11] Kevin James and Ethan Smith. Average Frobenius distribution for elliptic curves defined over finite Galois extensions of the rationals. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 150, pages 439–458. Cambridge Univ Press, 2011.

[JTT+16] Kevin James, Brandon Tran, Minh-Tam Trinh, Phil Wertheimer, and Dania Zantout. Extremal primes for elliptic curves. *Journal of Number Theory*, 164:282–298, 2016.

[Kou15] Dimitris Koukoulopoulos. Primes in short arithmetic progressions. *International Journal of Number Theory*, 11(05):1499–1521, 2015.

[Lou93] Stéphane Louboutin. Majorations explicites de $L|(1,\chi)|$. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 316(1):11–14, 1993.

[LT76] Serge Lang and Hale Trotter. *Frobenius Distributions in GL_2-Extensions: Distribution of Frobenius Automorphisms in GL_2-Extensions of the Rational Numbers*, volume 504. Springer, 1976.

[Par15] James Parks. Amicable pairs and aliquot cycles on average. *International Journal of Number Theory*, 11(6):1751–1790, 2015.

[Sil86] Joseph Silverman. *The Arithmetic of Elliptic Curves*, volume 106. 1986.

[Tay08] Richard Taylor. Automorphy for some ℓ-adic lifts of automorphic mod ℓ Galois representations. II. *Publications mathématiques*, 108(1):183–239, 2008.

Department of Mathematical Sciences, Clemson University, O-110 Martin Hall, Box 340975, Clemson, SC 29634

E-mail address: lgibers@g.clemson.edu

Department of Mathematical Sciences, Clemson University, O-110 Martin Hall, Box 340975, Clemson, SC 29634

E-mail address: kevja@clemson.edu

26 Oct 2017 17:44:55 PDT
Version 2 - Submitted to Acta Arith.