The Average Analytic Rank of Elliptic Curves

D.R. Heath-Brown
Mathematical Institute, Oxford

Abstract

All the results in this paper are conditional on the Riemann Hypothesis for the \( L \)-functions of elliptic curves. Under this assumption, we show that the average analytic rank of all elliptic curves over \( \mathbb{Q} \) is at most 2, thereby improving a result of Brumer [2]. We also show that the average within any family of quadratic twists is at most \( 3/2 \), improving a result of Goldfeld [3]. A third result concerns the density of curves with analytic rank at least \( R \), and shows that the proportion of such curves decreases faster than exponentially as \( R \) grows. The proofs depend on an analogue of Weil’s “explicit formula”.

1 Introduction

The purpose of this paper is to establish upper bounds for the average of the analytic rank of elliptic curves defined over \( \mathbb{Q} \). The article by Rubin and Silverberg [12] gives an excellent survey of this topic. Our first result concerns the average over all such curves, and sharpens an estimate of Brumer [2]. We introduce at the outset the minor technical trick of counting the curves

\[ E = E_{r,s} : y^2 = x^3 + rx + s \]

with a weight

\[ w_T(E) = w_1(T^{-1/3}r)w_2(T^{-1/2}s), \]

where \( w_1, w_2 \) are infinitely differentiable non-negative functions of compact support, vanishing at the origin. We define \( \Delta_E = -16(4r^3 + 27s^2) \) and we write

\[ \mathcal{C} = \{ E_{r,s} : p^4 | r \Rightarrow p^6 \; \text{divides} \; s, \; \Delta_E \neq 0 \} \]

and

\[ S(T) = \sum_{E \in \mathcal{C}} w_T(E). \]

Our principal result is then the following.

**Theorem 1** Assume that the \( L \)-functions of all the curves \( E_{r,s} \) satisfy the Riemann Hypothesis. Then

\[ \frac{1}{S(T)} \sum_{E \in \mathcal{C}} w_T(E)r(E) \leq 2 + o(1), \]

as \( T \to \infty \), where \( r(E) \) is the analytic rank of \( E \).
Thus the average analytic rank, taken over all elliptic curves defined over $\mathbb{Q}$, is at most 2. This improves on the corresponding result of Brumer [2], in which it was shown that the average is at most 2.3.

We shall also investigate the proportion of elliptic curves $E$ which have large rank. We define two sets,

$$D(T) = \{E_{r,s} : |r| \leq T^{1/3}, |s| \leq T^{1/2}, \Delta_{E_{r,s}} \neq 0\},$$

and

$$C(T) = \{E_{r,s} \in D(T) : p^4 | r \Rightarrow p^6 \nmid s\}.$$  

We then have the following result.

**Theorem 2** Assume that the $L$-functions of all the curves $E_{r,s}$ satisfy the Riemann Hypothesis. Then for any positive integer $R$ we have

$$\frac{\#\{E \in C(T) : r(E) \geq R\}}{\#C(T)} \ll (3R/2)^{-R/12},$$

where the implied constant is absolute.

Thus the proportion of curves with rank $R$ decreases faster than exponentially. We remark that it may be possible to improve the values of the constants $3/2$ and $12$ which occur in the theorem. We have merely given the simplest values that the method allows.

Since

$$\#C(T) \ll T^{5/6},$$

it follows that

$$r(E) \leq 11 \frac{\log T}{\log \log T}$$

for $E \in C(T)$ and sufficiently large $T$. Such results are already known (see Mestre [10] and Brumer [2]). However the fact that our theorem actually contains this estimate demonstrates that we have achieved the best rate of decay with respect to $R$ that one can currently hope for.

We shall also consider the set of quadratic twists

$$E_D : Dy^2 = x^3 + rx + s$$

of a fixed elliptic curve $E$ of conductor $N$, say. This family has previously been investigated by Goldfeld [3]. It is of some interest to separate the odd rank twists from those of even rank. We therefore define $L_D(s)$ to be the $L$-function of $E_D$, and $w_D = \pm 1$ to be the sign of the functional equation for $L_D(s)$. Thus if $w$ is the corresponding sign for the original curve $E$ we have

$$w_D = w \frac{D}{|D|} \chi_D(N), \quad (1.1)$$

for $(D,N) = 1$, where $\chi_D$ is the real primitive character associated to the quadratic field $\mathbb{Q}(\sqrt{D})$. We then set

$$T = \{D : (D,N) = 1\},$$

and

$$T^\pm = \{D \in T : w_D = \pm 1\},$$
where $D$ is restricted to run over fundamental discriminants in each case. It follows that $r(E_D)$ is even for $D \in T^+$, and odd for $D \in T^-$.

For technical reasons we find it convenient to count the twists $E_D$ with a smooth weight. We therefore choose a three times differentiable non-negative function $w(x)$, supported on a compact subset of either $(-\infty, 0)$ or $(0, \infty)$, and we define

$$W^\pm(T) = \sum_{D \in T^\pm} w(D/T).$$

Our result is then the following.

**Theorem 3** Let $E$ be a fixed elliptic curve defined over $\mathbb{Q}$. Suppose that the functions $L_D(s)$ all satisfy the Riemann Hypothesis. Then

$$\frac{1}{W^\pm(T)} \sum_{D \in T^\pm} w(D/T) r(E_D) \leq \frac{3}{2} + o(1),$$

as $T \to \infty$.

Of course it is natural to apply this result with a weight $w$ which approximates the characteristic function of an interval. Thus within a family of quadratic twists the average analytic rank would be at most $\frac{3}{4}$, whether one restricts to curves of odd rank or to curves of even rank. This may be compared with a result of Goldfeld [3, Proposition 2], who considers the set $T$ only, and in which the constant $\frac{3}{4}$ is replaced by $\frac{13}{4}$.

The reader should note that our theorem requires $L_D(s)$ to satisfy the Riemann Hypothesis for every integer $D$, even though the sets $T^\pm$ contain only integers $D$ which are coprime to $N$.

Naturally we expect that the above results should remain true if we replace the analytic rank $r(E)$ by the arithmetic rank, which we denote by $R(E)$. Results of Kolyvagin [8], [9] and Gross and Zagier [4] show that

$$R(E) = \begin{cases} 0, & \text{if } r(E) = 0, \\ 1, & \text{if } r(E) = 1. \end{cases}$$

(1.2)

Theorem 3 then has the following corollary.

**Theorem 4** Let $E$ be an elliptic curve defined over $\mathbb{Q}$. Suppose that the functions $L_D$ all satisfy the Riemann Hypothesis. Then

$$\frac{1}{W^+(T)} \sum_{D \in T^+, R(E_D)=0} w(D/T) \geq \frac{1}{W^+(T)} \sum_{D \in T^+, R(E_D)=0} w(D/T) \geq \frac{1}{4} + o(1)$$

and

$$\frac{1}{W^-(T)} \sum_{D \in T^-, R(E_D)=1} w(D/T) \geq \frac{1}{W^-(T)} \sum_{D \in T^-, R(E_D)=1} w(D/T) \geq \frac{3}{4} + o(1)$$

as $T \to \infty$. 

3
Thus at least 1/4 of all curves in $T^+$ would have arithmetic rank 0, and at least 3/4 of all curves in $T^-$ would have rank 1.

Assuming the Riemann Hypothesis for the $L$-function of the curve $E$, write

$$\Sigma(E, h, X) = \sum_{\gamma} \hat{h}(\gamma \log \frac{X}{2\pi}),$$

where $\frac{1}{2} + i\gamma$ runs over the non-trivial zeros of the $L$-function, and where $h(t)$ is a continuous function of a real variable supported in $[-1, 1]$. Here the Fourier transform is given by (2.2). With rather little extra effort one should be able to show in the context of Theorem 1 that

$$\frac{1}{S(T)} \sum_{E \in \mathcal{C}} w_T(E) \Sigma(E, h, X) = h(0) \frac{\log T}{\log X} + \frac{\hat{h}(0)}{2} + o(1)$$  \hspace{1cm} (1.3)

for $T^{\delta} \leq X \leq T^{2/3 - \delta}$, for any fixed $\delta > 0$. On choosing the function $h(t)$ so that $\hat{h}(t)$ is real and non-negative we would then get an estimate of the type given in Theorem 1, and the particular choice (2.1) leads to the upper bound $2 + o(1)$. It is the admissible range of $X$ which is of paramount importance here, since a larger range would permit an upper bound smaller than $2 + o(1)$. The relation (1.3) was established by Brumer [2, Theorem 4.4] for the shorter range $X \leq T^{5/9 - \delta}$, and this leads to a weaker version of our Theorem 1, with $2 + o(1)$ replaced by $2.3 + o(1)$.

In a completely analogous way one should be able to adapt the proof of Theorem 3 to show that

$$\frac{1}{W^T(T)} \sum_{D \in T^\pm} w(D/T) \Sigma(E_D, h, X) = 2h(0) \frac{\log T}{\log X} + \frac{\hat{h}(0)}{2} + o(1),$$  \hspace{1cm} (1.4)

for $T^{\delta} \leq X \leq T^{2 - \delta}$, for any fixed $\delta > 0$. Theorem 3 would then follow from a suitable choice of the function $h(t)$.

The asymptotic formulae (1.3) and (1.4) reflect the distribution of the low-lying zeros of the $L$-functions. There have been a number of investigations into such questions in related contexts, and in every case these give partial confirmation of the random matrix models for the distribution functions. The interested reader is referred to the survey by Katz and Sarnak [6] for an overview of the area, and to their book [7] for the background theory. Results closely analogous to (1.4) are given by Özlük and Snyder [11], who treat the zeros of quadratic $L$-functions, and by Iwaniec, Luo and Sarnak [5], whose work concerns zeros of families of automorphic forms.

Finally we should mention the work of Silverman [13], who considers general algebraic families of elliptic curves. His result is weaker than Theorem 3 when specialized to the family of quadratic twists, but is much more general. The interesting feature of Silverman’s result is that the upper bound depends on the rank of corresponding elliptic surface.

**Acknowledgements.** The bulk of this work was carried out in 1991, at the Institute for Advanced Study, Princeton. It was prepared for publication more recently while the author was a guest of the American Institute of Mathematics. It is a pleasure to thank both these bodies for their hospitality and financial assistance.
2 Preliminaries

Our starting point is the ‘explicit formula’ in the form given by Brumer [2, §2]. We apply this to an arbitrary elliptic curve $E$, which is of course now known to be modular, following the work of Wiles [15], Taylor and Wiles [14] and Breuil et. al. [1]. We write $N_E$ for the conductor of $E$, and

$$L_E(s) = \sum_{n=1}^{\infty} a_n(E)n^{-s}$$

for the $L$-function of $E$. We note that if $p \nmid N_E$ then

$$a_p(E) = \alpha_p + \overline{\alpha}_p$$

where

$$|\alpha_p| = \sqrt{p}.$$  

For the remaining primes $p|N_E$ the coefficients $a_p(E)$ are always 0, 1 or $-1$. Finally we define

$$c_{p^k}(E) = \begin{cases} -a_p(E)^k/p^k, & p|N_E, \\ -(\alpha_p^k + \overline{\alpha}_p^k)/p^k, & p \nmid N_E. \end{cases}$$

We take the weight function $F(t)$ in [2, Lemma 2.1] to be

$$F(t) = h_X(t) = h(t/\log X)$$

where $X \geq 2$ and

$$h(t) = \begin{cases} 1 - |t|, & |t| \leq 1, \\ 0, & |t| > 1. \end{cases} \quad (2.1)$$

We define the Fourier transform of a function $f(x)$ by

$$\hat{f}(t) = \int_{-\infty}^{\infty} e^{-2\pi i t x} f(x) dx. \quad (2.2)$$

Although this convention differs from Brumer’s, the two alternative definitions of $\hat{f}(0)$ agree. Since $h(0) = h(0) = 1$ and $h(t) \geq 0$ for all real $t$, we deduce from the estimates of Brumer [2, §2] that

$$r(E) \leq \frac{\log N_E}{\log X} + \frac{2}{\log X} (U_1(E, X) + U_2(E, X)) + O\left(\frac{1}{\log X}\right), \quad (2.3)$$

where

$$U_k(E, X) = \sum_{p^k \leq X, p \geq 5} c_{p^k}(E)(\log p^k)h_X(\log p^k).$$

In particular we have

$$U_1(E, X) = -\sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p)a_p(E).$$
We begin by considering \( U_2(E, X) \). For Theorem 1 we can use Brumer’s work [2, p. 457], which shows that

\[
\sum_{E \in \mathcal{C}} w_T(E)U_2(E, X) = \mathcal{S}(T)\left\{ \frac{h(0)}{4} \log X + O(T^{-1/2}X^{1/2} \log X) + O(T^{-3/4}X^{9/10} \log X)^{9/5} \right\} = \mathcal{S}(T)\left\{ \frac{1}{4} \log X + O(1) \right\}, \quad (2.4)
\]

providing that \( X \leq T^{5/6-\delta} \) with a fixed \( \delta > 0 \).

Similarly for Theorem 3 we may use the work of Goldfeld [3, p. 116] which produces

\[
U_2(E_D, X) = \frac{h(0)}{4} \log X + O(\log \log D) = \frac{1}{4} \log X + O(\log \log D), \quad (2.5)
\]

with an implied constant depending on \( E \). Finally, for Theorem 2 we note that

\[
|c_{p,k}(E)| \leq 2k^{-1}p^{-k/2},
\]

from the definition, whence

\[
|U_2(E, X)| \leq 2 \sum_{5 \leq p \leq \sqrt{X}} \frac{\log p}{p} \leq \log X + O(1). \quad (2.6)
\]

It remains therefore to consider the behaviour of \( U_1(E, X) \), for which we shall require slightly different techniques in each case.

Before leaving this section we need to record one further result given by Brumer [2, (2.13)]. In view of Brumer’s ‘Note added in proof’ [2, p. 472], we may state the result as follows.

**Lemma 1** Let \( k \) be an even \( C^1 \) continuous function with support in \([-1, 1]\), and suppose that \( \hat{k}(t) = O_\delta((1 + |t|)^{-1-\delta}) \) for some \( \delta > 0 \). Then if \( E \) is an elliptic curve of conductor \( N_E \) we have

\[
\sum_{p \leq X} \frac{\log p}{p} k\left( \frac{\log p}{\log X} \right) a_p(E) \ll \delta (\log N_E)(\log X)\{||k||_\infty + ||(1 + |t|)^{1+\delta}k||_\infty\}.
\]

Although Brumer proves this only when \( X > 10 \log N_E \) it is automatically true for smaller \( X \), by virtue of the bound \( |a_p(E)| \leq 2\sqrt{p} \).

### 3 Theorem 1—Initial Transformations

Let \( D \) be the set of all curves \( E_{r,s} \), including those for which \( \Delta = 0 \). For any curve \( E_{r,s} \in D \) we define

\[
\sigma_p(E_{r,s}) = -\tau_p^{-1} \sum_{t, x \pmod{p}} \left( \frac{t}{p} \right)^r e_p(tx^3 + txr + ts), \quad (3.1)
\]

where \( \tau_p \) is the usual Gauss sum. Here we adopt the standard convention that \( e_p(x) = \exp(2\pi ix/p) \). Then, according to Brumer [2, (3.1)], we have

\[
a_p(E) = \sigma_p(E) \quad (p \geq 5) \quad (3.2)
\]
for every $E \in \mathcal{C}(T)$. As Brumer remarks, this formula is valid even when $E$ is an elliptic curve with singular reduction modulo $p$. We also observe that for every $E_{r,s} \in \mathcal{D}$ we have

$$
\sigma_p = -\sum_{x \mod p} \left( \frac{x^3 + xr + s}{p} \right),
$$

whence

$$
|\sigma_p| \leq 2\sqrt{p}.
$$

(3.3)

The key estimate required for Theorem 1 is then as follows.

**Lemma 2** For any $\varepsilon > 0$ we have

$$
\sum_{p < p \leq 2p} |\sum_{E \in \mathcal{D}} w_T(E)\sigma_p(E)| \ll p^{\varepsilon}(p^{1/2}T^{5/6} + p^{3/2}T^{1/2} + p^{2}T^{1/6} + p^{7/2}T^{-1}).
$$

Here, and throughout the paper, we allow the constant implied by the $\ll$ symbol to depend on $\varepsilon$.

For convenience of notation we shall write

$$
\sum_{p < p \leq 2p} |\sum_{E \in \mathcal{D}} w_T(E)\sigma_p(E)| = \Sigma.
$$

We begin the proof of Lemma 2 by observing that the value $t = 0$ in (3.1) may be omitted, since $\left( \frac{0}{p} \right) = 0$. We can then substitute $y = tx$ for $x$, giving

$$
\sigma_p(E_{r,s}) = -\tau_p^{-1} \sum_{t \not\equiv 0 \mod p} \sum_{y \mod p} \left( \frac{t}{p} \right) e_p(t^{-2}y^3 + yr + ts).
$$

(Here we interpret $t^{-2}y^3$ modulo $p$.) Hence

$$
|\sum_{E \in \mathcal{D}} w_T(E)\sigma_p(E)| \leq p^{-1/2} |\sum_{t,y} \left( \frac{t}{p} \right) e_p(t^{-2}y^3)S_1S_2|,
$$

where

$$
S_1 = \sum_{r=-\infty}^{\infty} w_1(T^{-1/3}r)e_p(yr)
$$

and

$$
S_2 = \sum_{s=-\infty}^{\infty} w_2(T^{-1/3}s)e_p(ts).
$$

According to the Poisson summation formula the sum $S_1$, for example, is

$$
T^{1/3} \sum_{m=-\infty}^{\infty} \hat{w}_1(T^{1/3}m + \frac{y}{p}).
$$

Moreover, since $w_1$ has derivatives of all orders, it follows that

$$
\frac{d^n}{dx^n} \hat{w}_1(x) \ll_{n,A} (1 + |x|)^{-A}
$$

(3.4)
for any real $x$, any fixed integer $n \geq 0$, and any fixed $A > 0$. To bound the sum over $m$ it is convenient to fix the range for $y$ so that $|y| \leq p/2$. We then conclude that

$$\sum_{m=-\infty}^{\infty} \hat{w}_1(T^{1/3}(m + \frac{y}{p})) \ll 1,$$

and

$$\sum_{m \neq 0} \hat{w}_1(T^{1/3}(m + \frac{y}{p})) \ll T^{-A/3},$$

for $|y| \leq p/2$. The sum $S_2$ may be handled similarly, and we conclude that

$$S_1S_2 = T^{5/6}\hat{w}_1(-\frac{yT^{1/3}}{p})\hat{w}_2(-\frac{T^{1/2}}{p}) + O(T^{5/6-A/3}).$$

Moreover any terms for which $P/2 < |y| \leq p/2$ or $P/2 < |t| \leq p/2$ are $O(T^{5/6-A/3})$. We therefore arrive at the estimate

$$\Sigma \ll P^{-1/2}T^{5/6} \sum_p \left| \sum_{|t| \leq p/2} \hat{w}_2\left(-\frac{T^{1/2}}{p}\right)\frac{t}{p}\Sigma_1(t,p)\right| + P^{5/2}T^{5/6-A/3},$$

where

$$\Sigma_1(t,p) = \sum_{|y| \leq p/2} \hat{w}_1\left(-\frac{yT^{1/3}}{p}\right)e_p(t^{-2}y^3).$$

If we take $A = 6$ the final term is

$$P^{5/2}T^{-7/6} \leq P^{7/2}T^{-1},$$

which is satisfactory for Lemma 2.

It may be worth observing at this point that the bound (3.4), together with its analogue for $\hat{w}_2$, yields

$$\sum_{t,p} |\hat{w}_1\left(-\frac{yT^{1/3}}{p}\right)e_p(t^{-2}y^3)| \ll (1 + T^{-1/3}P)(1 + T^{-1/2}P),$$

whence one trivially has

$$\Sigma \ll P^{1/2}(T^{1/3} + P)(T^{1/2} + P) + P^{7/2}T^{-1}.$$
and strictly positive on $[1, 2]$. We then have

$$\sum_p \left| \sum_{|t| \leq P/2} w_2 \left( \frac{T^{1/2}}{p} \right) \left( \frac{t}{p} \right) \sum_1(t, p) \right| \leq \frac{P}{2} t \neq 0 \hat{w}_2 \left( \frac{T^{1/2}}{k} \right) \left( \frac{t}{k} \right) \sum_1(t, k).$$

Here we define the Jacobi symbol $(t/k)$ to be zero whenever $k$ is even. We now wish to bring the summation over $t$ outside the modulus signs. In order to do this we observe that

$$\max_{P/2 \leq k \leq 5P/2} \left| \hat{w}_2 \left( \frac{xP}{k} \right) \right| \ll A (1 + |x|)^{-A}$$

for any $A > 0$. Hence, on defining

$$M(t) = \min \{ 1, (|t| T^{1/2} P^{-1})^{-1} \},$$

we see that

$$\max_{P/2 \leq k \leq 5P/2} \left| \hat{w}_2 \left( \frac{T^{1/2}}{k} \right) \right| \ll A M(t)^A,$$

for any $A > 0$. We now have

$$\sum_k w_3(kP^{-1}) \sum_{0 < |t| \leq P/2} \hat{w}_2 \left( \frac{T^{1/2}}{k} \right) \left( \frac{t}{k} \right) \sum_1(t, k) \ll \sum_k w_3(kP^{-1}) \sum_{t \neq 0} M(t)^A |\sum_1(t, k)|,$$

whence

$$\Sigma \ll P^{-1/2} T^{5/6} \sum_k w_3(kP^{-1}) \sum_{(t,k)=1} M(t)^A |\sum_1(t, k)| + P^{7/2} T^{-1}. \quad (3.6)$$

4 Lemma 2—The Kernel of the Proof

In order to perform the averaging over $k$ we shall use Cauchy’s inequality to reverse the order of summations in (3.6). In view of (3.5) and the fact that $w_3(x)$ is supported on $[\frac{1}{2}, \frac{5}{2}]$, we have

$$\sum_{k,t} w_3(kP^{-1}) |t|^{-1} M(t)^A \ll P \log P.$$
On applying Cauchy’s inequality to (3.6) we deduce that

$$\sum \ll P^{-1/2}T^{5/6}\{P \log P\}^{1/2}\Sigma_1^{1/2} + P^{7/2}T^{-1},$$

(4.1)

where

$$\Sigma_1 = \sum_k w_3(kP^{-1}) \sum_{(t,k)=1} |t|M(t)^A|\Sigma_1(t,k)|^2$$

$$= \sum_t |t|M(t)^A \sum_{|y_1|,|y_2| \leq P/2} \sum_{(k,t)=1} w_4(kP^{-1}) e_k(t^{-2}\{y_1^3 - y_2^3\}).$$

with

$$w_4(x) = w_3(x)\tilde{w}^3(x)\frac{y_1T^{1/3}}{P^x}\tilde{w}^1(x)\frac{y_2T^{1/3}}{P^x}.$$ 

Thus $w_4$ is supported in $[\frac{1}{2}, \frac{5}{2}]$, and for any fixed integer $n \geq 0$ and any $A > 0$ we have

$$\frac{d^n}{dx^n}w_4(x) \ll_{A,n} m(y_1)^A m(y_2)^A,$$

(4.2)

by (3.4), where

$$m(y) = \min\{1, (\frac{|y|T^{1/3}}{P^{-1}})^{-1}\}. $$

(4.3)

We now write

$$a = y_1^3 - y_2^3, \quad b = t^2.$$ 

If $b\bar{b} \equiv 1(\text{mod} \, k)$ and $k\bar{k} \equiv 1(\text{mod} \, b)$, then $b\bar{b} + k\bar{k} \equiv 1(\text{mod} \, bk)$, for $(b,k) = 1$, so that

$$\frac{\bar{b}}{k} + \frac{k}{\bar{b}} = \frac{1}{bk}$$

is an integer. Thus $e_k(a\bar{b}) = e_{bk}(a)e_b(-a\bar{k})$. With this in mind we define

$$\rho(x) = w_4(x)e\left(\frac{a}{bP^x}\right),$$

so that

$$\Sigma_1 = \sum_{t|y_1,y_2} |t|M(t)^A \sum_{(k,t)=1} \rho(kP^{-1}) e_k(-a\bar{k}).$$

We decompose the inner sum into residue classes modulo $b$, and apply the Poisson summation formula to obtain

$$\sum_{j \equiv j \equiv 1(\text{mod} \, b)} e_b(-aj) \sum_{k \equiv j \equiv 1(\text{mod} \, b)} \rho(kP^{-1}) e_b(-a\bar{k})$$

$$= \sum_j e_b(-aj) \sum_{m=-\infty}^\infty e_b(nj) \frac{P}{b} \hat{\rho}\left(\frac{nP}{b}\right).$$

(4.4)

At this point we observe that

$$\hat{\rho}(x) \ll x^{-2} \sup_v |\rho''(v)|$$
and that
\[ \rho''(x) \ll (1 + \left| \frac{a}{bP} \right|^2)m(y_1)^A m(y_2)^A \]
by (4.2), whence
\[ \rho(\frac{nP}{b}) \ll n^{-2}\left( \frac{b^2}{P^2} + \frac{a^2}{P^4} \right)m(y_1)^A m(y_2)^A. \]

However
\[ \frac{b^2}{P^2} + \frac{a^2}{P^4} \ll P^2T^{-2}\left\{ \left( \frac{|t|T^{1/2}}{P} \right)^4 + \left( \frac{|y_1|T^{1/3}}{P} \right)^6 + \left( \frac{|y_2|T^{1/3}}{P} \right)^6 \right\} \]
by (3.5) and (4.3). It therefore follows that the terms \( n \neq 0 \) in (4.4) are
\[ \ll \sum_{j \pmod{b}} \sum_{|t| = 1} P^{-2} m(t)^{-4} m(y_1)^{A-6} m(y_2)^{A-6} \]
\[ \ll P^3T^{-2}M(t)^{-4} m(y_1)^{A-6} m(y_2)^{A-6}. \]

On choosing \( A = 8 \) we see that the contribution to \( \Sigma_1 \) is
\[ P^3T^{-2} \sum_{t,y_1,y_2} \left| t \right| \left| M(t) \right|^4 m(y_1)^2 m(y_2)^2 \]
\[ \ll P^3T^{-2} \left\{ \sum_{t} \left| t \right| \left| M(t) \right|^4 \right\} \left\{ \sum_{y} m(y)^2 \right\}^2 \]
\[ \ll P^3T^{-2}(1 + PT^{-1/2})^2(1 + PT^{-1/3})^2 \]
by (3.5) and (4.3). The contribution to \( \Sigma \) itself is then
\[ \ll P^{3/2}T^{-1/6}(\log P)^{1/2}(1 + PT^{-1/2})(1 + PT^{-1/3}) \]
\[ \ll P^\varepsilon \left( P^{3/2}T^{-1/6} + P^{5/2}T^{-1/2} + P^{7/2}T^{-1} \right), \]
by (4.1), and this is satisfactory for Lemma 2, since
\[ P^{5/2}T^{-1/2} = \left\{ P^{3/2}T^{1/2}, P^{7/2}T^{-1} \right\}^{1/2} \]
\[ \leq \max \left\{ P^{3/2}T^{1/2}, P^{7/2}T^{-1} \right\} \]
\[ \leq P^{3/2}T^{1/2} + P^{7/2}T^{-1}. \]

5 Lemma 2—A Highest Common Factor Sum

It remains to handle the terms \( n = 0 \) in (4.4). Since
\[ |\rho(0)| \leq \int_{-\infty}^{\infty} |\rho(x)| dx = \int_{-\infty}^{\infty} |w_4(x)| dx \ll m(y_1)^A m(y_2)^A, \]
by (4.2), the contribution to \( \Sigma_1 \) is
\[ \ll \sum_{t,y_1,y_2} \frac{P}{b} \left| t \right| m(t)^A m(y_1)^A m(y_2)^A \left| \sum_{j \pmod{b}} e_b(-a) \right|. \quad (5.1) \]
The sum over \( j \) is a Ramanujan sum which may be evaluated as
\[
\sum_{d | (a, b)} d \mu(b/d).
\]
We therefore see that (5.1) is
\[
\ll P \sum_{t=1}^{\infty} t^{\varepsilon-1} M(t)^A \sum_{|y_1| \leq y_2} m(y_2)^A(a, b),
\]
for any \( \varepsilon > 0 \), where \( (a, b) \) denotes the highest common factor of \( a \) and \( b \). The terms with \( y_1 = y_2 = 0 \) are
\[
P \sum_{t=1}^{\infty} t^{\varepsilon+1} M(t)^A \ll P^{3+\varepsilon} T^{-1} + P^{1+\varepsilon}.
\]
For the remaining sum we shall use the following lemma.

Lemma 3 For any \( U, V \geq 1 \) and any \( \varepsilon > 0 \), we have
\[
\sum_{1 \leq u \leq U} \sum_{|w| \leq V} (u^2, v^3 - w^3) \ll U^{1+\varepsilon} V(U^2 + V).
\]
We now see that the ranges \( U/2 < |t| \leq U \) and \( V/2 < y_2 \leq V \) contribute
\[
\ll PU^{\varepsilon-1} M(U)^A m(V)^A U^{1+\varepsilon} V(U^2 + V)
\]
to (5.2), and hence to \( \Sigma_1 \). We choose \( A = 4 \), say, and sum \( U \) and \( V \) over powers of 2 to obtain a total contribution to \( \Sigma_1 \) of
\[
\ll P^{3+\varepsilon} T^{-1} + P^{1+2\varepsilon}(1 + P^{3T-4/3} + P^2 T^{-2/3}).
\]
This is satisfactory for Lemma 2, by (4.1).

It remains to prove Lemma 3. We write \( S \) for the sum to be estimated, and we take
\[
(u^2, v^3 - w^3) = d = \prod p^e.
\]
We also define
\[
\delta = \prod p^{[(e+1)/2]},
\]
where \([x]\) denotes the integer part of \( x \), as usual. It follows that \( \delta | u \) whenever \( d | u^2 \), so that \( u \) takes at most \( U/\delta \) values for each given value of \( d \). We therefore see that
\[
S \leq U \sum_{d \leq U^2} d \delta^{-1} \# \{ v, w : d \mid v^3 - w^3 \}.
\]
We now consider the value of \( (d, v^3) \) which we denote by \( \alpha = \prod p^{f} \). On defining
\[
\beta = \prod p^{[(f+2)/3]},
\]
we see that \( \beta | v \), and since \( (d, v^3) = (d, w^3) \) we also must have \( \beta | w \). We may therefore write \( v = \beta v' \) and \( w = \beta w' \), whence
\[
v^3 \equiv w^3 \pmod{\gamma},
\]
(5.3)
where \( \gamma = \prod p^\varrho \) with
\[
g = \max\{e - 3\left[\frac{f + 2}{3}\right], 0\}.
\]
By construction we have \((v', \gamma) = 1\) so that the congruence \((5.3)\) has at most \(3^w(\gamma) \ll U^\varepsilon\) solutions \(w' \pmod{\gamma}\), for each value of \(v'\). (Here \(\omega(\gamma)\) is the number of distinct prime factors of \(\gamma\).) Thus, for given values of \(d, \alpha, \beta\) and \(\gamma\), there are at most \(1 + \frac{V}{\beta}\) possible choices of \(v\), to each of which there correspond \(O\left(U^{1+\varepsilon}\right)\) possible values of \(w\). We therefore conclude that
\[
S \ll U^{1+\varepsilon} \sum_{d, \alpha, \beta, \gamma} \frac{d}{\delta} \left(1 + \frac{V}{\beta} + \frac{V^2}{\beta^2 \gamma}\right)
\]
\[
\ll U^{1+\varepsilon} \sum_{d, \alpha, \beta, \gamma} \frac{d}{\delta} \left(V + \frac{V^2}{\beta^2 \gamma}\right).
\]

We now set
\[
f(d) = \prod_p p^{[e/2]},
\]
whence
\[
f(d) = \frac{d}{\delta},
\]
and
\[
g(d) = \prod_p p^{[e/3]-[(e+1)/2]},
\]
for which we claim that
\[
g(d) \geq \frac{d}{\delta \beta^2 \gamma}.
\]
To prove the latter it is enough to verify that
\[
\left[\frac{e}{3}\right] - \left[\frac{e+1}{2}\right] \geq e - \left[\frac{e+1}{2}\right] - 2\left[\frac{f + 2}{3}\right] - \max\{e - 3\left[\frac{f + 2}{3}\right], 0\}
\]
for \(0 \leq f \leq e\), which is an easy exercise. Moreover, since \(\alpha, \beta\) and \(\gamma\) all divide \(d\), they take \(O(U^{\varepsilon})\) values each. It follows that
\[
S \ll U^{1+4\varepsilon} \sum_{d \leq U^2} \{f(d)V + g(d)V^2\}.
\]
We now observe that the Dirichlet series \(\sum f(d)d^{-\sigma}\) and \(\sum g(d)d^{-\sigma}\) are convergent for \(\sigma > 1\) and \(\sigma > 0\) respectively, since their Euler products converge. We may therefore deduce that
\[
\sum_{d \leq U^2} f(d)d^{-1-\varepsilon} \ll_\varepsilon 1,
\]
whence
\[
\sum_{d \leq U^2} f(d) \leq \sum_{d \leq U^2} f(d)\left(\frac{U^2}{d}\right)^{1+\varepsilon} \ll_\varepsilon U^{2+2\varepsilon}.
\]
In a similar manner we find that
\[
\sum_{d \leq U^2} g(d) \ll_\varepsilon U^{2\varepsilon}.
\]
These bounds suffice for the proof of the lemma, on replacing \(\varepsilon\) by \(\varepsilon/6\).
6 Theorem 1—Completion of the Proof

Whenever \( w_T(E) \neq 0 \) we have \( \Delta_E \ll T \), and hence \( N_E \ll T \). It therefore follows from (2.3) and (2.4) that

\[
\frac{1}{S(T)} \sum_{E \in C} w_T(E) r(E) \leq \frac{\log T}{\log X} + \frac{1}{2} + O\left(\frac{1 + U_1}{\log X}\right)
\]

for \( X \leq T^{2/3} \), where

\[
U_1 = \frac{1}{S(T)} \left| \sum_{p \leq X} \frac{\log p}{p} h_X(\log p) \sum_{E \in C} w_T(E) a_p(E) \right|,
\]

by (3.2). We proceed to show that if \( \delta > 0 \) is fixed, then \( U_1 \ll 1 \) when \( X = T^{2/3-\delta} \). This suffices for Theorem 1.

We begin by considering the contribution made by the set of singular curves, which we denote by \( S \). If \( E \) is singular then \( \Delta_E = 0 \), whence

\[
\sum_{E \in S} w_T(E) \ll T^{1/6}.
\]

Moreover one may verify, using the definition (3.1), that \( |\sigma_p(E)| \leq 1 \) for \( E \in S \), whence

\[
\sum_{P < p \leq 2P} \left| \sum_{E \in S} w_T(E) \sigma_p(E) \right| \ll PT^{1/6}.
\]

(6.1)

For the non-singular curves \( E = E_{r,s} \) we put \( r = d^4 \rho \) and \( s = d^6 \sigma \), where \( d \) is a positive integer and \( E_{\rho,\sigma} \in C \). For each curve \( E = E_{r,s} \) we write \( E^* \) for the corresponding curve \( E_{\rho,\sigma} \), so that

\[
w_T(E) = w_{T_{d^{-12}}}(E^*).
\]

Moreover \( \sigma_p(E) = \sigma_p(E^*) \), if \( p \mid d \), and \( \sigma_p(E) = 0 \) otherwise. It follows that

\[
\sum_{E \in \mathcal{D} - S} w_T(E) \sigma_p(E) = \sum_{d=1}^{\infty} \sum_{\rho \in \mathcal{D}} w_{T_{d^{-12}}}(E) \sigma_p(E) + \theta(T, p),
\]

where

\[
\theta(T, p) = - \sum_{d \equiv 0 (\text{mod } p)} \sum_{E \in \mathcal{C}} w_T(d^{-12} E) \sigma_p(E) \ll T^{5/6} p^{-10} p^{1/2}
\]

by the bound (3.3). The Möbius inversion formula now yields

\[
\sum_{E \in \mathcal{C}} w_T(E) \sigma_p(E) = \sum_{d=1}^{\infty} \mu(d) \sum_{E \in \mathcal{D} - S} w_{T_{d^{-12}}}(E) \sigma_p(E) = \sum_{d=1}^{\infty} \mu(d) \theta(Td^{-12}, p),
\]

and the second sum on the right is

\[
\ll \sum_{d=1}^{\infty} (Td^{-12})^{5/6} p^{-19/2} \ll T^{5/6} p^{-19/2}.
\]
We therefore see that
\[
\sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sum_{E \in \mathcal{C}} w_T(E) \sigma_p(E)
\]
\[=
\sum_{d=1}^{\infty} \mu(d) \sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sum_{E \in \mathcal{D}-S} w_{Td-12}(E) \sigma_p(E)
+ O\left( \sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) T^{5/6} p^{-10/2} \right)
\]
\[=
\sum_{d=1}^{\infty} \mu(d) \sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sum_{E \in \mathcal{D}-S} w_{Td-12}(E) \sigma_p(E)
+ O(T^{5/6}). \tag{6.2}
\]

Since \(S(T) \gg T^{5/6}\) the error term is satisfactory for the desired bound \(U_1 \ll 1\).

If \(d \gg T^{1/12}\) then \(w_{Td-12}(E)\) will vanish. Thus we may restrict the sum over \(d\) to the interval \(d \ll T^{1/12}\). We split this range at \(d = d_0\), with a value of \(d_0\) to be specified in due course, see (6.4).

When \(d \geq d_0\) we write
\[
\sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sum_{E \in \mathcal{D}-S} w_{Td}(E) \sigma_p(E)
\]
\[=
\sum_{E \in \mathcal{D}-S} w_{Td}(E) \sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sigma_p(E). \tag{6.3}
\]

To estimate the inner sum we put \(E = E_{r,s}\) and we let \(r = f^4 \rho\) and \(s = f^6 \sigma\) with \(E_{\rho,\sigma} \in \mathcal{C}\). For convenience we set \(E^* = E_{\rho,\sigma}^*\) as before. Thus \(\sigma_p(E) = 0\) for \(p | f\) and
\[\sigma_p(E) = \sigma_p(E^*) = a_p(E^*)\]
otherwise. It follows that
\[
\sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) \sigma_p(E) = \sum_{5 \leq p \leq X} \frac{\log p}{p} h_X(\log p) a_p(E^*)
- \sum_{5 \leq p \leq X, p | f} \frac{\log p}{p} h_X(\log p) a_p(E^*).
\]

The first sum on the right is \(O(\log T)^2\) by Lemma 1, while the second is trivially \(O(f)\), since \(|a_p(E^*)| \leq 2\sqrt{p}\). Thus the first sum contributes
\[
\ll \sum_{E \in \mathcal{D}-S} w_{Td-12}(E)(\log T)^2 \ll T^{5/6} d^{-10}(\log T)^2
\]
to (6.3), and the second
\[ \ll \sum_{r} \sum_{s} w_1 \left( \frac{r}{(Td^{-12})^{1/3}} \right) w_2 \left( \frac{s}{(Td^{-12})^{1/2}} \right) \sum_{f^{s}|r, f^{s}|s} f \]
\[ = \sum_{f=1}^{\infty} \sum_{r=0}^{\left( \text{mod } f^{s} \right)} \sum_{s=0}^{\left( \text{mod } f^{s} \right)} w_1 \left( \frac{r}{(Td^{-12})^{1/3}} \right) w_2 \left( \frac{s}{(Td^{-12})^{1/2}} \right) \]
\[ \ll \sum_{f=1}^{\infty} \frac{T^{5/6} d^{-10}}{f^{10}} \]
\[ \ll T^{5/6} d^{-10}. \]

We therefore see that
\[ \sum_{5 \leq p \leq X} \frac{\log p}{p} \mu_X(\log p) \sum_{E \in \mathbb{D} - S} w\left(T d^{-12}\right) \sigma_p(E) \ll T^{5/6} d^{-10} (\log T)^2. \]

Thus terms with \( d \geq d_0 \) contribute \( O(T^{5/6} d_0^{-9} (\log T)^2) \) to (6.2). On choosing
\[ d_0 = \log T, \] say, we see that this is \( O(T^{5/6}) \), which is satisfactory.

For the values \( d < d_0 \) we observe that
\[ \sum_{P < p \leq 2P} \left| \sum_{E \in \mathbb{D} - S} w\left(T d^{-12}\right) \sigma_p(E) \right| \]
\[ \leq \sum_{P < p \leq 2P} \left| \sum_{E \in \mathbb{D}} w\left(T d^{-12}\right) \sigma_p(E) \right| + O(P(T d^{-12})^{1/6}), \] (6.5)
by a second application of (6.1). According to Lemma 2 the inner sum is
\[ \ll P^c (P^{1/2} T^{5/6} d^{-10} + P^{3/2} T^{1/2} d^{-6} + P^2 T^{1/6} d^{-2} + P^7/2 T^{-1} d^{12}), \]
whence
\[ \sum_{P < p \leq 2P} \left| \sum_{E \in \mathbb{D} - S} w\left(T d^{-12}\right) \sigma_p(E) \right| \]
\[ \ll P^c (P^{1/2} T^{5/6} d^{-10} + P^{3/2} T^{1/2} d^{-6} + P^2 T^{1/6} d^{-2} + P^7/2 T^{-1} d^{12}), \]

since the error term \( O(PT^{1/6} d^{-2}) \) in (6.5) is majorized by the term
\[ P^c, P^2 T^{1/6} d^{-2} \]
above. It follows that
\[ \sum_{5 \leq p \leq X} \frac{\log p}{p} \mu_X(\log p) \sum_{E \in \mathbb{D} - S} w\left(T d^{-12}\right) \sigma_p(E) \]
\[ \ll T^{5/6} d^{-10} + X^{2c} (X^{1/2} T^{1/2} d^{-6} + X T^{1/6} d^{-2} + X^{5/2} T^{-1} d^{12}), \]
whence the terms with \( d < d_0 \) contribute
\[ \ll T^{5/6} + X^{2c} T^c (X^{1/2} T^{1/2} + X T^{1/6} + X^{5/2} T^{-1}) \]
to (6.2). If we choose \( X = T^{2/3 - \delta} \), and take \( \varepsilon \) sufficiently small in terms of \( \delta \), all these terms will be \( O(T^{5/6}) \). This is also satisfactory for the desired bound
\[ U_1 \ll 1. \] The proof of Theorem 1 is therefore complete.
7 Proof of Theorem 2

To establish Theorem 2 we combine (2.3) with the estimate (2.6) to show that

\[ r(E) \leq \frac{\log N_E}{\log X} + \frac{2}{\log X} U_1(E, X) + 2 + O\left(\frac{1}{\log X}\right). \]

It will be convenient to remove the first few primes from the sum \( U_1(E, X) \), so we shall write

\[ U_1(E, X) = U(E, X) + O(1), \]

where

\[ U(E, X) = \sum_{100 < p \leq X} c_p(E) h_X (\log p) \log p, \]

say. Then, since the curves under consideration have \( N_E \ll T \), we deduce that

\[ r(E) \leq 2 + \frac{\log T}{\log X} + \frac{2}{\log X} U(E, X) + O\left(\frac{1}{\log X}\right). \]

Consequently, if \( X \geq X_0 \), where \( X_0 \) is a sufficiently large absolute constant, and if

\[ r(E) \geq R \geq 3 + 2 \frac{\log T}{\log X}, \tag{7.1} \]

then

\[ |U(E, X)| \geq \frac{1}{2} \log T. \]

We complete the proof of the theorem by estimating moments of the sum \( U(E, X) \). Under the hypothesis (7.1) we see that

\[ \# \{ E \in C(T) : r(E) \geq R \} \left(\frac{1}{2} \log T\right)^{2k} \leq \sum_{E \in C(T)} |U(E, X)|^{2k}, \]

for any positive integer \( k \). We now set

\[ V(E, X) = \sum_{100 < p \leq X} \frac{\log p}{p} h_X (\log p) \sigma_p(E) \]

for any \( E \in D(T) \), so that \( U(E, X) = V(E, X) \) whenever \( E \in C(T) \), by (3.2). We then have

\[ \# \{ E \in C(T) : r(E) \geq R \} \left(\frac{1}{2} \log T\right)^{2k} \leq \sum_{E \in D(T)} |V(E, X)|^{2k}, \tag{7.2} \]

for any positive integer \( k \). We note that \( V(E, X) \) is in fact real, and expand \( |V(E, X)|^{2k} \) by the multinomial theorem. This gives

\[ \sum_{E \in D(T)} \sum_{e} C(e) F(e), \tag{7.3} \]

where

\[ F(e) = \prod_{100 < p \leq X} \{ \log p h_X (\log p) \sigma_p(E) \}^{e_p}. \]
Here \( \mathbf{e} \) runs over vectors with one non-negative integer component \( e_p \) for each prime \( p \in (100, X] \), and such that \( \sum e_p = 2k \). Moreover the multinomial coefficients \( C(\mathbf{e}) \) are given by

\[
C(\mathbf{e}) = \frac{(2k)!}{\prod e_p!}.
\]

We divide the terms in (7.3) into two classes. Type I terms will be those for which every exponent \( e_p \) satisfies either \( e_p = 0 \) or \( e_p \geq 2 \). The remaining terms will be type II terms.

We begin by considering type I terms. Since \( |\sigma_p(E)| \leq 2p^{-1/2} \) by (3.3), we have

\[
|F(\mathbf{e})| \leq \prod_p \left( \frac{2h_X \log p \log p}{\sqrt{p}} \right)^{f_p},
\]

where

\[
f_p = \begin{cases} 0, & e_p = 0, \\ 2, & e_p \geq 2. \end{cases}
\]

Here we use the fact that

\[
\frac{2h_X \log p \log p}{\sqrt{p}} \leq 1
\]

for \( p > 100 \). Moreover \( C(\mathbf{e}) \leq (2k)! \) for every vector \( \mathbf{e} \). Thus the terms for which exactly \( j \) primes have \( f_p = 2 \) can contribute at most

\[
\frac{(2k)!}{j!} S^j
\]

to (7.2), where

\[
S = \sum_{100 < p \leq X} \frac{(2h_X \log p \log p)^2}{p}.
\]

We now observe that, with our choice of \( h_X \), we have

\[
S = \frac{\log^2 X}{3} + O(\log X).
\]

Thus, if \( k \leq \log X \) with \( X \) sufficiently large, the contribution to (7.2) from all type I terms is at most

\[
\ll T^{5/6} \sum_{j \leq k} \frac{(2k)!}{j!} S^j \ll T^{5/6} \frac{(2k)!}{k!} \left( \frac{\log^2 X}{3} \right)^k \ll T^{5/6} \left( \frac{4k \log^2 X}{3e} \right)^k.
\]  

(7.4)

We turn now to the type II terms. We begin by recalling the definition

\[
\sigma_p(E_{r,s}) = -\tau_p^{-1} \sum_{t,x \mod p} \left( \frac{t}{p} \right) e_p(t r^3 + t x r + ts).
\]

When we sum over \( E \in \mathcal{D}(T) \) we have therefore to estimate

\[
\sum_{E_{r,s} \in \mathcal{D}(T)} \sum_{t_1, \ldots, t_{2k}} \sum_{x_1, \ldots, x_{2k}} \prod_{i=1}^{2k} \left( \frac{t_i}{p_i} \right) e_{p_i} (t_i x_i^3 + t_i x_i r + t_i s),
\]

18
where \( t_i \) and \( x_i \) run modulo \( p_i \), and the primes \( p_1, \ldots, p_{2k} \) include at least one value, \( p^* \) say, which is not repeated. We bound the above expression as

\[
\ll \left( \prod_{i} p_i \right) \sum_{r} \sum_{t_1, \ldots, t_{2k}} \left| \sum_{s} e(s \{ t_1/p_1 + \ldots + t_{2k}/p_{2k} \} ) \right|
\]

in which \( t_i \) runs over \( 1, \ldots, p_i - 1 \), and the function \( e(x) \) is given by \( e(x) = \exp(2\pi ix) \). The summation conditions on \( r \) and \( s \) are given by

\[
|r| \leq T^{1/3}, \quad |s| \leq T^{1/2}, \quad 4r^3 + 27s^2 \neq 0.
\]

We proceed to examine the innermost sum. We write

\[
\frac{t_1}{p_1} + \ldots + \frac{t_{2k}}{p_{2k}} = \alpha
\]

and note that \( \alpha \) cannot be an integer, since its denominator must be divisible by \( p^* \). It follows that

\[
||\alpha|| \geq \left( \prod_{i} p_i \right)^{-1},
\]

whence

\[
\left| \sum_s e(sa) \right| \leq 2 + \frac{1}{|\sin \pi \alpha|} \ll \frac{1}{||\alpha||} \leq \prod_{i} p_i.
\]

Here we have allowed for the fact that, for a given value of \( r \), the variable \( s \) runs over all integers in the interval \([-T^{1/2}, T^{1/2}]\) with at most 2 exceptions. Since \( r \) takes \( O(T^{1/3}) \) values we therefore see that

\[
\sum_{E \in \mathcal{D}(T)} F(e) \ll T^{1/3} \prod_{100 < p \leq X} \left\{ \frac{p^2 \log p}{p^{[\tau_p]}} \right\}^{e_p}
\]

for each type II term, whence the total contribution to (7.2) is

\[
\ll T^{1/3} \left( \sum_{100 < p \leq X} p^{1/2} \log p \right)^{2k} \ll T^{1/3} X^{3k}. \quad (7.5)
\]

In view of (7.2) and the estimates (7.4) and (7.5) we find that

\[
\# \{ E \in \mathcal{C}(T) : r(E) \geq R \} \left( \frac{1}{2} \log T \right)^{2k} \ll T^{5/6} \left( \frac{4k \log^2 X}{3e} \right)^k + T^{1/3} X^{3k}
\]

for \( X \geq X_0 \), subject to the conditions

\[
R \geq 3 + 2 \log \frac{T}{\log X}
\]

and \( k \leq \log X \). Note here that \( X_0 \) is independent of \( k \). We therefore choose

\[
X = T^{1/6k},
\]

whence

\[
\# \{ E \in \mathcal{C}(T) : r(E) \geq R \} \ll (27ek/4)^{-k} T^{5/6},
\]
for $R \geq 3 + 12k$ and $T^{1/6k} \geq X_0$. We take $k = \left\lfloor \frac{R-3}{12} \right\rfloor$ and write $j = R/12$, so that $k \leq j \leq k + O(1)$. Then for any positive constants $a > b$ we will have $(ak)^{-k} \ll (bj)^{-j}$ if $k$ is large enough. Since $27e/4 > 18$ we conclude that

$$\#\{E \in C(T) : r(E) \geq R\} \ll (3R/2)^{-R/12}T^{5/6}, \quad (7.6)$$

if $R$ is large enough and $R \ll \log T$. However (7.6) is trivially true for bounded values of $R$. Moreover for

$$R = \left\lfloor 11\frac{\log T}{\log \log T} \right\rfloor$$

we may already conclude from (7.6) that

$$\#\{E \in C(T) : r(E) \geq R\} = o(1),$$

so that there can be no curves with

$$r(E) \geq 11\frac{\log T}{\log \log T}$$

for large enough $T$. This completes the proof of Theorem 2.

8 Theorem 3—Preliminary Sieving

The condition $D \in T^\pm$ is distinctly awkward to work with, and our first task is therefore to replace it with something more manageable. When $N$ is odd we begin by decomposing $T^\pm$ according to the power of 2 dividing $D$. Of course, if $N$ is even then $D$ is automatically odd. We now write $D = \delta \hat{n}$ with $\delta = +1$ or $-1$ and $\hat{n}$ odd, and we decompose $T^\pm$ further according to the residue class of $n$ modulo 8. This produces a collection of triples $(k, \delta, e)$, in which

$$k = 1, 3, 5 \text{ or } 7, \quad \delta = +1 \text{ or } -1, \quad \text{and } e = 0, 2, \text{ or } 3,$$

and such that $T^\pm$ is a disjoint union of certain of the sets

$$\{D = \delta 2^n : w_D = \pm 1, \mu^2(n) = 1, n \equiv k \pmod{8}\}.$$ 

We shall prove the analogue of Theorem 3 for these sets, assuming that the weight function $w$ is supported on a compact subset of $(-\infty, 0)$ for $\delta = -1$, and $(0, \infty)$ for $\delta = +1$. Theorem 3 itself will then follow. Henceforth we shall regard the triple $(k, \delta, e)$ and the sign $\pm 1$ as fixed, and for any positive odd integer $n$ we shall write $D = D(n) = \delta \hat{n}$, where $\hat{n}$ is the square-free kernel of $n$. We also set

$$W(n/T) = w(\delta \hat{n}/T),$$

$$\mathcal{F} = \{n \in \mathbb{N} : w_D = \pm 1, (n, N) = 1, \mu^2(n) = 1, n \equiv k \pmod{8}\},$$

and

$$\mathcal{R}(T) = \sum_{n \in \mathcal{F}} W(n/T).$$

We have therefore to show that

$$\frac{1}{\mathcal{R}(T)} \sum_{n \in \mathcal{F}} W(n/T) r(E_D) \leq \frac{3}{2} + o(1),$$

20
as $T \to \infty$.

We turn now to the condition that $n$ must be square-free. We define

$$P = \prod_{2 < p \leq \log \log T} p,$$

and we set

$$X(n) = \sum_{d | P, d^2 | n} \mu(d),$$

so that $X(n) = 0$ if $n$ is divisible by the square of a prime $p \leq \log \log T$, and $X(n) = 1$ otherwise. It follows that

$$\sum_{n \in \mathcal{F}} W(n/T)r(E_D) \leq \sum_{n \in \mathcal{G}} X(n)W(n/T)r(E_D),$$

where

$$\mathcal{G} = \{ n \in \mathbb{N} : w_D = \pm 1, (n, N) = 1, n \equiv k \pmod{8} \}.$$

Moreover it is a straightforward matter to demonstrate the asymptotic formula

$$\sum_{n \in \mathcal{G}} X(n)W(n/T) \sim R(T),$$

since $R(T) \gg T$ if $\mathcal{F}$ is non-empty. It therefore suffices to establish the estimate

$$\sum_{n \in \mathcal{G}} X(n)W(n/T)r(E_D) \leq (\frac{3}{2} + o(1)) \sum_{n \in \mathcal{G}} X(n)W(n/T). \quad (8.1)$$

The proof of Theorem 3 now hinges on the following lemma.

**Lemma 4** Let $2 \leq X \leq T^{2-\varepsilon}$, where $\varepsilon$ is a positive constant. Then, assuming the Riemann Hypothesis for all the $L$-functions $L_D(s)$, we have

$$\sum_{n \in \mathcal{G}} W(n/T)U_1(E_{D(n)}, X) \ll T.$$

We conclude this section by demonstrating how (8.1) may be deduced from Lemma 4. We have

$$\sum_{n \in \mathcal{G}} X(n)W(n/T)U_1(E_D, X)$$

$$= \sum_{d | P} \mu(d) \sum_{n \in \mathcal{G}, d^2 | n} W(n/T)U_1(E_{D(n)}, X)$$

$$= \sum_{d | P} \mu(d) \sum_{m \in \mathcal{G}} W(d^2m/T)U_1(E_{D(m)}, X), \quad (8.2)$$

on replacing $n$ by $d^2m$. (Notice that $D(d^2m) = D(m)$, and that $m \in \mathcal{G}$ if and only if $d^2m \in \mathcal{G}$.) However

$$d \leq P = \exp\{O(\log \log T)\} \ll T^\varepsilon.$$
Thus if $X \leq T^{2-5\varepsilon}$ we have

$$X \leq \left(\frac{T}{d^2}\right)^{2-\varepsilon},$$

for any $\varepsilon > 0$. We may therefore apply Lemma 4 to the inner sum in (8.2), giving

$$\sum_{n \in \mathcal{G}} X(n) W(n/T) U_1(E_D, X) \ll \sum_{d \mid p} T \ll T$$

for $X = T^{2-5\varepsilon}$. We now feed (2.3) and (2.5) into the left-hand side of (8.1) and observe that

$$N_{E_D} \ll D^2 \ll T^2$$

since the curve $E$ is fixed. This produces the required bound (8.1). Thus to complete the proof of Theorem 3 it will suffice to establish Lemma 4.

### 9 Further Simplifications

In this section we shall simplify the expression occurring in Lemma 4. We begin by noting that

$$a_p(E_D) = (\frac{D}{p})a_p(E)$$

for primes $p \nmid ND$, and we therefore classify the odd primes $p$ according to their residue modulo 8, which enables us to write

$$(\frac{D}{p}) = \eta_h(\frac{n}{p}), \quad \text{for} \quad p \equiv h \pmod{8}, \quad p \nmid n,$$

where $\eta_h$ may depend on $\varepsilon, \delta$ and $e$ as well as on $h$. Thus

$$U_1(E_D, X) = -\sum_h \eta_h \sum_{p \equiv h \pmod{8}} \beta_p \left(\frac{n}{p}\right) + O\left(\sum_{p \mid n} \frac{\log p}{\sqrt{p}}\right),$$

where we have introduced the shorthand

$$\beta_p = \begin{cases} \frac{\log p}{\sqrt{p}} h_X (\log p) a_p(E), & p \geq 5, \\ 0, & p = 2, 3. \end{cases}$$

Since $E$ is fixed we have

$$\sum_{p \mid N} \frac{\log p}{\sqrt{p}} \ll 1,$$

so that these terms contribute $O(T)$ in Lemma 4. Moreover

$$\sum_n W(n/T) \sum_{p \mid n} \frac{\log p}{\sqrt{p}} \ll \sum_p \frac{\log p T}{\sqrt{p} p} \ll T.$$
which is also satisfactory. The condition \( p \equiv h \pmod{8} \) may be picked out by using an appropriate combination of the characters

\[
\left( \frac{a}{p} \right), \quad a = 1, -1, 2, -2.
\]

For the proof of Lemma 4 it therefore suffices to show that

\[
\sum_{n \in \mathcal{G}} W(n/T)U(an) \ll T,
\]

for \( a = 1, -1, 2, -2 \), where

\[
U(m) = \sum_p \beta_p \left( \frac{m}{p} \right).
\]

We turn now to the condition \( n \in \mathcal{G} \). Since \((n, N) = 1\), the root number \( w_D \) differs from \( (\frac{N}{n}) \) by a factor depending on \( N, k, \delta \) and \( \epsilon \) only, in view of (1.1). We can therefore pick out the conditions \( w_D = \pm 1 \) and \( n \equiv k \pmod{8} \) by introducing a suitable combination of factors \( (\frac{N}{n}), (\frac{2}{n}) \) and \( (\frac{-1}{n}) \). We deduce that it is sufficient, for the proof of Lemma 4, to establish the estimate

\[
\sum_{(n, 2N) = 1} \psi(n)W(n/T)U(an) \ll T,
\]

where \( \psi(n) \) is a real primitive character of conductor dividing \( 8N \). Since

\[
\sum_{(n, 2N) = 1} \psi(n)W(n/T)U(an) = \sum_d \mu(d) \sum_{d|n} \psi(n)W(n/T)U(an)
\]

\[
= \sum_d \mu(d) \psi(d) \sum_{m=1}^{\infty} \psi(m)W(dm/T)U(adm),
\]

we conclude as follows.

**Lemma 5** In order to establish Lemma 4 it suffices to show that

\[
\sum_{n=1}^{\infty} \psi(n)W(n/T)U(rn) \ll_{r,E} T,
\]

for \( 2 \leq X \leq T^{2-\varepsilon} \), and for every \( r \neq 0 \).

## 10 Character Sums

We now have to examine

\[
\sum_{n=1}^{\infty} \psi(n)W(n/T)U(rn) = \sum_p \beta_p \left( \frac{r}{p} \right) \sum_n W(n/T)\psi_p(n), \quad (10.1)
\]

where

\[
\psi_p(n) = \psi(n)\left( \frac{n}{p} \right). \quad (10.2)
\]
We shall denote the sum on the left of (10.1) by $\Sigma$. The primes for which $p \mid N$, contribute a total $O(T)$ to $\Sigma$. For the remaining primes $\psi_p$ is primitive. We write $\Sigma_p$ for the inner sum on the right of (10.1), and we denote the conductor of $\psi_p$ by $q$. Thus $q = bp$, say, where $b$ is the conductor of $\psi$. Moreover $b \mid 8N$. We proceed to decompose $\Sigma_p$ by dividing the values of $n$ into congruence classes $n \equiv j \pmod{q}$, whence

$$\Sigma_p = \sum_{j \pmod{q}} \psi_p(j) \sum_{m=-\infty}^{\infty} W\left(\frac{j + qm}{T}\right).$$

On applying the Poisson summation formula we obtain

$$\Sigma_p = \sum_{j \pmod{q}} \psi_p(j) \sum_{m=-\infty}^{\infty} \frac{T}{q} e\left(\frac{mj}{q}\right) \hat{W}\left(\frac{Tm}{q}\right)$$

$$= \frac{T}{q} \sum_{m=-\infty}^{\infty} \hat{W}\left(\frac{Tm}{q}\right) \sum_{j \pmod{q}} \psi_p(j) e\left(\frac{mj}{q}\right),$$

where $e(x) = \exp\{2\pi ix\}$ as usual. On writing $G(p)$ for the Gauss sum

$$\sum_{j \pmod{q}} \psi_p(j) e\left(\frac{j}{q}\right),$$

we have

$$\sum_{j \pmod{q}} \psi_p(j) e\left(\frac{mj}{q}\right) = G(p) \psi_p(m),$$

so that

$$\Sigma_p = \frac{T G(p)}{q} \sum_{m=-\infty}^{\infty} \hat{W}\left(\frac{Tm}{q}\right) \psi_p(m).$$

Since $\psi_p(0) = 0$, we therefore conclude that

$$\Sigma = \frac{T}{b} \sum_{m \neq 0 \atop m \not\mid N} \frac{G(p)}{p} \beta_p W\left(\frac{Tm/b}{p}\right) \psi_p(m)\left(\frac{r}{p}\right) + O(T). \quad (10.3)$$

It is instructive to examine the trivial estimate for $\Sigma$ at this stage. Since the function $W$ is supported on a compact subset of $(0, \infty)$, and is three times differentiable, we have

$$\hat{W}(x) \ll \min\{1, |x|^{-3}\}. \quad (10.4)$$

Thus, on using the bounds $G(p) \ll p^{1/2}$ and $\beta_p \ll p^{-1/2} \log p$, we find that

$$\Sigma \ll T + T \sum_{m \neq 0 \atop m \not\mid X} \frac{\log p}{p} \frac{p^3}{T^3|m|^3} \ll T + T^{-2} X^3.$$

This therefore suffices for an analogue of Lemma 5 in which $X$ may be as large as $T^{1-\epsilon}$. One would then obtain a version of Theorem 3 with a constant $\frac{5}{2}$ in place of $\frac{3}{2}$. Such an improvement of Goldfeld’s bound was mentioned by Brumer.
[2, p. 445], although it is clear that the argument intended by Brumer was a relatively minor modification of that used by Goldfeld.

Our sharper estimate for $\Sigma$ stems from a non-trivial bound for the inner sum in (10.3). To obtain this we call on the following ‘Prime Number Theorem’ for twisted curves $E_D$.

**Lemma 6** If $L_D(s)$ satisfies the Riemann Hypothesis we have

$$\sum_{p \leq x} \frac{a_p(E)}{p} \chi_D(p) \log p \ll x^\varepsilon |D|^\varepsilon$$

for any $\varepsilon > 0$, where the implied constant depends at most on $E$ and $\varepsilon$.

Here the reader should recall that $\chi_D$ is the real primitive character associated to the quadratic field $\mathbb{Q}(\sqrt{D})$. (When $D = 1$ we take $\chi_D$ to be the trivial character.) We shall prove Lemma 6 in the next section. Notice that the lemma does not assume that $D$ and $N$ are coprime.

To apply Lemma 6 to (10.3) we observe that

$$G(p) = \psi(p)\left(\frac{b}{p}\right)\tau(p)\overline{\tau(p)} = C_b\psi'(p)\sqrt{p}(1 - i(\frac{-1}{p})),$$

by the usual evaluation of Gauss sums. Here $C_b$ is a constant depending on $b$ only, and $\psi'$ is a real character whose modulus divides $8N$. In view of the definition (10.2) of $\psi_p$, it follows that there is a real character $\psi_1$ whose modulus divides $8Nmr$, such that

$$\sum_{p \mid N} G(p)\beta_p \overline{W}\left(\frac{Tm/b}{p}\right)\psi_p(m)\left(\frac{r}{p}\right) \ll \left| \sum_{p \mid 30Nr} \frac{a_p(E)}{p^{3/2}}(\log p)\log p \overline{W}\left(\frac{Tm/b}{p}\right)\psi_1(p) \right|.$$

We now wish to replace $\psi_1$ by the primitive character $\chi_\Delta$ which induces it. Here $\Delta$ is a fundamental discriminant and $\Delta | 8Nmr$. This process will introduce an error which contributes

$$\ll T \sum_{m \neq 0} \sum_{p \mid 30Nmr} \frac{\log p}{p} \frac{p^3}{T^3|m|^3}$$

to $\Sigma$, by (10.4). The primes dividing $30Nr$ provide at most

$$\ll T \sum_{m \neq 0} \sum_{p \mid 30N_r} \frac{\log p}{p} \frac{p^3}{T^3|m|^3} \ll T^{-2},$$

and the primes $p|m$ yield a total

$$\ll T \sum_{m \neq 0} \sum_{p|m, p \leq X} \frac{\log p}{p} \frac{p^3}{T^3|m|^3} \ll T^{-2} \sum_{p \leq X} \frac{p^2 \log p}{p} \sum_{m \neq 0, p|m} |m|^{-3} \ll T^{-2} \sum_{p \leq X} \frac{p^2 (\log p)}{p^{-3}} \ll T^{-2} \log X.$$
Both these contributions are satisfactory, and we conclude from (10.3) that
\[ \Sigma \ll T + T \sum_{m \neq 0} \left| \sum_p \frac{a_p(E)}{p^{3/2}} \log p \right| \chi_X \left( \frac{\log p}{p} \right) \hat{W} \left( \frac{Tm/b}{p} \right) \chi_{m}(p). \] (10.5)

We shall bound the sum over \( p \) by using partial summation together with the estimate
\[ \sum_{p \leq x} \frac{a_p(E)}{p} \log p \ll x^\varepsilon |m|^{\varepsilon} \]
which follows from Lemma 6. In analogy to (10.4) we have
\[ \frac{d}{dx} \hat{W}(x) \ll \min\{1, |x|^{-3}\}. \]

We deduce that
\[
\int_2^X \left| \frac{d}{dt} \left\{ t^{-1/2} h_X(\log t) \hat{W} \left( \frac{Tm/b}{t} \right) \right\} \right| dt \\
\ll \int_2^X t^{-3/2} \min\{1, (\frac{T|m|}{t})^{-1} \} dt \\
\ll (T|m|)^{-1/2} \min\{1, (\frac{X}{T|m|})^{3/2} \}.
\]

We therefore conclude, on summing by parts, that
\[
\sum_p \frac{a_p(E)}{p^{3/2}} \log p \chi_X \left( \frac{\log p}{p} \right) \hat{W} \left( \frac{Tm/b}{p} \right) \chi_{m}(p) \\
\ll (X|m|)^\varepsilon (T|m|)^{-1/2} \min\{1, (\frac{X}{T|m|})^{3/2} \}.
\]

In view of (10.5) we now have
\[
\Sigma \ll T + T^{1/2} X^\varepsilon \sum_{m \neq 0} |m|^{-1/2 + \varepsilon} \min\{1, (\frac{X}{T|m|})^{3/2} \} \\
\ll T + X^{1/2 + 2\varepsilon},
\]
from which Lemma 5 follows on redefining \( \varepsilon \).

11 Proof of Lemma 6

To prove Lemma 6 we shall apply Lemma 1 to the curve \( E_D \), taking the function \( k \) to be
\[ k(t) = \frac{X h(t) - (X - 1) h(\frac{t}{X-1})}{\log^2 X}. \]

Since
\[ \hat{k}(t) = \frac{X}{\log^2 X} \frac{\sin^2(\pi t) - \sin^2(\pi(1 - X^{-1})t)}{\pi^2 t^2} \]

26
the hypothesis of Lemma 1 is satisfied for any \( \delta > 0 \). Moreover one readily finds that \( \|k\|_\infty \ll \frac{1}{(\log X)^2} \) and that

\[
\hat{k}(t) \ll \frac{X}{t^2(\log x)^2} \min\{1, |t|\} \min\{1, \frac{|t|}{X}\},
\]

whence

\[
\left\| (1 + |t|)^{1+\delta} \hat{k}(t) \right\|_\infty \ll \frac{X^\delta}{(\log X)^2}.
\]

It follows that

\[
\sum_{p \leq X} \frac{\log p}{p} k\left( \frac{\log p}{\log X} \right) a_p(E_D) \ll X^\delta \log |D|.
\]

However

\[
k\left( \frac{\log p}{\log X} \right) = \begin{cases} 
(\log X)^{-2}, & p \leq X^{1-1/X}, \\
O((\log X)^{-2}), & X^{1-1/X} \leq p \leq X,
\end{cases}
\]

whence

\[
\sum_{p \leq X} \frac{\log p}{p} k\left( \frac{\log p}{\log X} \right) a_p(E_D)
= (\log X)^{-2} \sum_{p \leq X} \frac{\log p}{p} a_p(E_D) + O(X^{-1/2} \log^2 X).
\]

It therefore follows, on choosing \( \delta = \varepsilon/2 \), that

\[
\sum_{p \leq X} \frac{\log p}{p} a_p(E_D) \ll X^\varepsilon |D|^{\varepsilon}.
\]

To complete the proof of Lemma 6 it remains to observe that the only primes for which \( a_p(E_D) \) can differ from \( a_p(E_D)^{\chi_D(p)} \) are, possibly, those for which \( p|30ND \). Since \( N \) is fixed, these contribute \( O(|D|^{\varepsilon}) \), which is satisfactory.

12 Deduction of Theorem 4

To prove Theorem 4 we begin by observing that

\[
\sum_{D \in T^+, r(E_D)=0} w(D/T) \geq \sum_{D \in T^+, r(E_D)=0} w(D/T),
\]

by (1.2). Moreover, Theorem 3 yields

\[
\sum_{D \in T^+ \times r(E_D)=0} w(D/T) = \sum_{D \in T^+} w(D/T) - \sum_{D \in T^+, r(E_D) \geq 2} w(D/T) \\
\geq W^+(T) - \sum_{D \in T^+, r(E_D) \geq 2} w(D/T) \frac{r(E_D)}{2} \\
= W^+(T) - \sum_{D \in T^+} w(D/T) \frac{r(E_D)}{2} \\
\geq W^+(T) - \frac{3}{2} \left[ \frac{3}{2} + o(1) \right] W^+(T) \\
= \left( \frac{1}{4} + o(1) \right) W^+(T),
\]
since \( r(E_D)/2 \geq 1 \) whenever \( r(E_D) \geq 2 \).

Similarly we have

\[
\sum_{D \in T^-, r(E_D) = 1} w(D/T) \geq \sum_{D \in T^-, r(E_D) = 1} w(D/T),
\]

and

\[
\sum_{D \in T^-, r(E_D) = 1} w(D/T) = \sum_{D \in T^-} w(D/T) - \sum_{D \in T^-, r(E_D) \geq 3} w(D/T)
\geq W^-(T) - \sum_{D \in T^-, r(E_D) \geq 3} w(D/T) \frac{r(E_D) - 1}{2}
= W^-(T) - \sum_{D \in T^-} w(D/T) \frac{r(E_D) - 1}{2}
\geq \frac{3}{2} W^-(T) - \frac{1}{2} \left( \frac{3}{2} + o(1) \right) W^-(T)
= \left( \frac{3}{4} + o(1) \right) W^-(T),
\]
as required for Theorem 4.

References


Mathematical Institute,
24-29, St. Giles’,
Oxford
OX1 3LB
England

rhb@maths.ox.ac.uk