Orders of CM elliptic curves modulo $p$ with at most two primes.

Henryk Iwaniec and Jorge Jiménez Urroz

1 Introduction and statement of results.

Nowadays the generation of cryptosystems requires two main aspects. First the security, and then the size of the keys involved in the construction and communication process. About the former one needs a difficult mathematical assumption which ensures your system will not be broken unless a well known difficult problem is solved. In this context one of the most famous assumption underlying a wide variety of cryptosystems is the computation of logarithms in finite fields and the Diffie Hellman assumption. However it is also well known that elliptic curves provide good examples of representation of abelian groups reducing the size of keys needed to guarantee the same level of security as in the finite field case. The first thing one needs to perform elliptic logarithms which are computationally secure is to fix a finite field, $\mathbb{F}_p$, and one curve, $E/\mathbb{F}_p$ defined over the field, such that $|E(\mathbb{F}_p)|$ has a prime factor as large as possible.

In practice the problem of finding such a pair, of curve and field, seems simple, just take a curve with integer coefficients and a prime $p$ of good reduction at random and see if $|E(\mathbb{F}_p)|$ has a big prime factor. However the theory that makes the previous algorithm useful is by no means obvious, neither clear or complete. For example it is well known that supersingular elliptic curves have to be avoided in the previous process since they reduce the security of any cryptosystem based on the Diffie Hellman assumption on the elliptic logarithm. But more importantly, the process will be feasible whenever the probability to find a pair, $(E, p)$, with a big prime factor $q | |E(\mathbb{F}_p)|$ is big enough. One problem arises naturally from the above.

Problem 1 Let $x$ be a positive number, $E/\mathbb{Q}$ an elliptic curve over the rational, and consider the sequence $\hat{\mathcal{A}}(x) = \{|E(\mathbb{F}_p)| : p \leq x\}$. How many elements $a \in \hat{\mathcal{A}}(x)$ have a large prime factor?

Before start let us note that, since the reduction modulo $p$ injects $E(\mathbb{Q})_{\text{tors}}$ into $E(\mathbb{F}_p)$ for almost all primes $p$, if the curve has rational torsion all the elements in $\hat{\mathcal{A}}(x)$ have a trivial common factor. In this sense, if $d$ is a common factor of all the elements in $\mathcal{A}(x)$, we will be considering the more convenient sequence $\mathcal{A}(x) = \{ \frac{1}{d}|E(\mathbb{F}_p)| : p \leq x\}$. 

1
The problem then would be to study the factorization into prime numbers of the elements of the sequence $A = A(\infty)$. This sequence has been widely study in the literature. In 1988 Koblitz [K] conjectured that for any elliptic curve over the rationals, the elements in $A$ not only have a big prime factor very frequently, but in fact there are infinitely many of them being prime numbers. Concretely if we denote by $\Pi_E(x)$ the function which counts the number of $a \in A$, $a \leq x$, that are primes, then he claims that for curves without rational torsion there exist a constant $c$, depending on the curve, such that $\Pi_E(x) \sim cx/(\log x)^2$ as $x \to \infty$. But there is also another reason why one would like to know the factorization of the elements in $A$. In 1977 Lang and Trotter conjectured that, given an elliptic curve $E$ and a nontorsion point $a \in E(\mathbb{Q})$, the density of primes for which $a$ generates $E(F_p)$ exists. In particular it predicts that the group of $F_p$ points of the reduced curve mod $p$ is cyclic for many primes $p$. Since then it has been an extensive study, either of the conjecture itself, or in the cyclicity of the group of $F_p$ points. A few examples can be founded in [B-M-P], [C1], [G-M], [L-T] or [S].

Both problems, to find lower bounds for the prime factors, and to ensure cyclicity of the group can be studied at the same time. In particular if we are able to prove that many elements in $A$ are squarefree, then automatically, at least when $d = 1$, the corresponding group will be cyclic. But if we are able to also say that the number of its prime factors is small, then one of them has to be big in comparison with the size of the element. Hence, to attack both problems, we want to find squarefree elements in $A$ with small number of prime factors. We are in a good position to understand our question as part of a general framework, namely inside sieve theory. Let us say that an integer $n$ is $P_r$ if it is squarefree with at most $r$ prime factors. If $r = 2$ we say our number is almost prime. In these terms we are interested in localizing many $P_r$ among the elements of $A$ with as small $r$ as possible. In particular Koblitz’ conjecture deals with the case $r = 1$.

Although the most efficient techniques known at present to attack this kind of problems are sieve methods, however it is important to note that, unfortunately, at least considered in its classical way, the sieve can not provide us with lower bounds for the number of primes in certain sequences due to the parity problem. In fact when $r = 1$ there is not a single example of a curve for which the asymptotics predicted by Koblitz have been proved. For $r > 1$ the situation is not much more promising although now, with sieve methods, one can accomplish something. Miri and Murty in [M-M] proved, assuming the Grand Riemann Hypothesis, GRH, that for curves without complex multiplication $|\{P_6 \in A(x)\}| \gg x/(\log x)^2$. In [S-W] Steuding and Weng improved the previous result giving $|\{P_6 \in A(x)\}| \gg x/(\log x)^2$ for non-CM curves. They also proved $|\{P_6 \in A(x)\}| \gg x/(\log x)^2$ in the CM case, but always under GRH. Very recently Cojocaru in [C2] proved unconditionally that for CM elliptic curves $|\{P_6 \in A(x)\}| \gg x/(\log x)^2$. 

2
In this paper we will also be considering curves with complex multiplication, focusing in the non-supersingular case. We shall improve the previous works unconditionally. For simplicity we will restrict our arguments to the curve

\[ E : y^2 = x^3 - x, \]

although the general CM case could be treated in a similar way. Note that any elliptic curve over \( \mathbb{Q} \) can only have complex multiplication by an order of an imaginary quadratic field of class number one. In our case the ordinary primes are \( p \equiv 1 (\mod 4) \) and, for these, \( |E(\mathbb{F}_p)| = p + 1 - 2a \) where \( p = a^2 + b^2 \), and \( a + ib \equiv 1 (\mod 2(i + i)) \) and so we deduce that 8 always divides \( |E(\mathbb{F}_p)| = (a - 1)^2 + b^2 \).

**Theorem 2** For \( x \geq 5 \) we have

\[ \left| \{ p \leq x , p \equiv 1 (\mod 4) : \frac{1}{2} |E(\mathbb{F}_p)| = P_2 \} \right| \gg \frac{x}{(\log x)^2}. \]

It is important to note that for these primes \( p \) in Theorem 2 of size about \( x \) one of the prime divisors of the \( P_2 \) has to be of order at least \( \sqrt{x} \).

Several remarks might be needed. First of all we see that we have supersingular reduction for any prime \( p \equiv 3 (\mod 4) \). Although this case might be of less interest for cryptographic purposes, it is interesting to see what happens. Now \( |E(\mathbb{F}_p)| = p + 1 \) is always divisible by \( d = 4 \) and, if we ask whether or not the elements of \( \mathcal{A} \) can be prime, we will be asking for primes \( q \) such that \( p + 1 = 4q \) for some prime \( q \), and this is well known to be essentially equivalent to the twin prime conjecture so the best one can hope for is the analogous results of Chen [Ch] for this problem. The case \( p \equiv 1 (\mod 4) \) is also related with the twin prime conjecture. Indeed, we have already mentioned that in this case \( p = a^2 + b^2 \) for some integers \( a, b \) which, looking at the problem in the gaussian domain, is just saying that \( p \) splits in \( \mathbb{Z}[i] \) as \( p = \pi \bar{\pi} \) for \( \pi = a + bi \). If we take \( \pi \) to be primary, i.e. \( \pi \equiv 1 (\mod 2(1 + i)) \), then \( |E(\mathbb{F}_p)| = N(\pi - 1) \equiv 0 (\mod 8) \), and so the elements of the sequence \( \mathcal{A} \) will be prime if and only if there exists a prime \( \hat{\pi} \) such that \( \pi - 1 = 2(1 + i)\hat{\pi} \), for some unit \( \epsilon = \pm 1, \pm i \). In other words, the problem for \( p \equiv 1 (\mod 4) \) is equivalent to the twin prime conjecture, but in the gaussian domain. For the proof of Theorem 2 we will apply techniques similar to those by Chen, but in the domain of gaussian integers. Among other things we will need to employ the so called switching principle, and also to extend the Bombieri-Vinogradov theorem in two different ways. The first generalization needed is the analogous result for the field \( \mathbb{Q}(i) \). Among many generalizations that occur in the literature in this direction, we appeal to [J], which is suitable to our particular case. The second generalization is a Bombieri-Vinogradov type theorem, not for primes, but rather for gaussian \( P_3 \) type numbers.

2 **A weighted sum for the sieve problem.**

Let us introduce the notation we will need afterwards. As usual for any sequence of rational integers, \( C \), and positive number \( x \), we will denote \( C(x) = \{ c \in C : \)
\[ c \leq x \}, \text{ and } |C(x)| \text{ the number of elements in the set. Given an integer } d, \text{ the set } \\
\{ c \in C : d \mid c \} \text{ consists of the elements of } C \text{ which are multiples of } d \text{ and } \\
S(C, d) = |\{ c \in C : (c, d) = 1 \}| \text{ counts the number of elements in } C \text{ coprime with } d. \text{ Analogously we define } \\
\xi(C, \delta) \text{ for } \mathbb{C} \subset \mathbb{Z}[i] \text{ and } \delta \in \mathbb{Z}[i]. \text{ We will also make several useful conventions. From now on } \lambda, \lambda_1, \lambda_2, \ldots, \text{ denote primes in } \mathbb{Z}[i] \text{ and } l, l_1, l_2, \ldots \text{ the rational primes below them. Furthermore } p, p_0, p_1, p_2, p_3 \text{ will be rational primes splitting in } \mathbb{Z}[i], \text{ and } \pi, \pi_0, \pi_1, \pi_2, \pi_3 \text{ will denote primary gaussian primes above them. On the other hand } q \text{ will be a rational prime inert in the domain. We put } \\
P(z) = \left\{ p \equiv 1 \pmod{4} : p \leq z \right\}, \text{ and } P(z) = \prod_{p \in P(z)} p, \text{ and on the other hand, } \\
Q(z) = \left\{ q \equiv 3 \pmod{4} : q \leq z \right\}, \text{ and } Q(z) = \prod_{q \in Q(z)} q. \text{ In order to prove Theorem 2 we first translate the problem in terms of gaussian integers. Let } \\
A(x) = \left\{ a = N \left( \frac{x-1}{2(1+i)} \right), \ |\pi|^2 \leq x \right\}, \text{ and } \\
S(x) = \sum_{p_2 \in A(x)} 1. \text{ Then it is clear that } S(x) \text{ is twice the left hand side of the inequality in Theorem 2. Therefore it suffices to prove } \\
S(x) \gg x/(\log x)^2. \tag{1} \\
A weighted sum will be considered to achieve this goal. In particular let } \\
W(x) = \sum_{a \in A(x)} \left\{ 1 - \sum_{p_0 \leq y} \frac{1}{2} - \sum_{z < p_0 \leq y} \frac{1}{2} \right\} \\
= \sum_{a \in A(x)} 1 - \frac{1}{2} \sum_{z < p_0 \leq y} - \sum_{a = p_1 p_2 p_3} \frac{1}{2} \\
= W_1(x) - \frac{1}{2} W_2(x) - \frac{1}{2} W_3(x), \tag{2} \\
\text{where } z = x^{1/8} \text{ and } y = x^{1/3}. \text{ It is clear that any term with positive weight in } W(x) \text{ has to be either } P_2 \text{ or divisible by some nontrivial square. however, the contribution to } W(x) \text{ of non-squarefree elements are easily bounded by } \\
\sum_{p > z} \sum_{n \leq x \mod p^2} \tau(n) \ll \frac{x \log x}{z \log z}.
Hence
\[ S(x) \geq \sum_{(P_2, zP(z)) = 1} 1 \geq W(x) + O\left(\frac{x \log x}{z \log z}\right), \]
and so, in order to prove the theorem we need the estimation
\[ W(x) \gg x/(\log x)^2. \]
We will estimate \( W_1(x), W_2(x), W_3(x) \) separately.

3 Lower bound for \( W_1(x) \).

First let us note that \( W_1(x) = S(A(x), 2P(z)Q(z)) \) is the usual sum in sieve theory which counts the elements in the sequence \( A(x) \) coprime with a product of certain primes, in this case \( 2P(z)Q(z) \). In order to estimate this sum we need to have some control on \( A_d(x) = \{ a \in A(x) : d\mid a \} \) for any \( d\mid 2P(z)Q(z) \). We will write \( d \) as \( d = 2^{d_1}d_2 \) such that \( d_1\mid P(z) \) and \( d_2\mid Q(z) \). For that purpose we will use the following

Lemma 3 Let \( C \) be a sequence of integers. For \( x > 0 \) and \( d \) squarefree we have
\[ |C_d(x)| = \sum_{k\mid d} \mu(k)S(C(x), k). \]

Proof: It follows by Möbius inversion formula.

Now it is clear that for any squarefree integer \( k \), and \( \alpha \in \mathbb{Z}[i] \) we have \( (N(\alpha), k) = 1 \) if and only if \( (\alpha, \kappa) = 1 \) where \( \kappa = k \) if \( k \) is odd and \( \kappa = (1+i)k/2 \) when \( k \) is even. Hence, \( S(A(x), k) = S(\mathfrak{A}(x), \kappa) \), where
\[ \mathfrak{A}(x) = \left\{ \frac{\pi - 1}{2(1+i)} : |\pi| < x \right\}, \]
and so, by Lemma 3,
\[ |A_d(x)| = \sum_{k\mid d} \mu(k)S(\mathfrak{A}(x), \kappa). \quad (3) \]

In order to estimate \( S(\mathfrak{A}(x), \kappa) \) we will use the inclusion-exclusion principle over the ideals in \( \mathbb{Z}[i] \). In particular let us define the Möbius function
\[ \hat{\mu}(\mathfrak{d}) = \begin{cases} 1 & \text{if } \mathfrak{d} = \langle 1 \rangle \\ (-1)^r & \text{if } \mathfrak{d} = \lambda_1 \cdots \lambda_r, \lambda_i \text{ distinct,} \\ 0 & \text{if } \lambda^2 \mid \mathfrak{d}. \end{cases} \]

It is easy to see that \( \hat{\mu}(\cdot) \) is a multiplicative function over the ideals in \( \mathbb{Z}[i] \) which verifies
\[ \sum_{\mathfrak{d}\mid \alpha} \hat{\mu}(\mathfrak{d}) = \begin{cases} 1 & \alpha = \langle 1 \rangle \\ 0 & \text{otherwise,} \end{cases} \]
and so for any $\kappa \in \mathbb{Z}[i]$ we have
\[ S(\mathfrak{A}(x), \kappa) = \sum_{\alpha \in \mathfrak{A}(x)} \sum_{d \mid \langle \alpha, \kappa \rangle} \hat{\mu}(d) = \sum_{d \mid \kappa} \hat{\mu}(d)|\mathfrak{A}(x)|. \quad (4) \]

Hence, the problem reduces to compute $|\mathfrak{A}(x)|$ for ideals $\mathfrak{d} | (1 + i)P(z)Q(z)$. By definition we have
\[ |\mathfrak{A}(x)| = \Pi(x; 2|1 + i|) = \left\{ \begin{array}{ll} \Pi(x; 2, 1) & \text{if } (1 + i) \nmid \mathfrak{d}, \\ \frac{1}{2} \Pi(x; 2, 1) + R_{\mathfrak{d}}(x) & \text{if } (1 + i) \mid \mathfrak{d} \end{array} \right. \]
where, for a general ideal $\mathfrak{a} \in \mathbb{Z}[i]$, and gaussian integer $\alpha$, we write
\[ \Pi(x; \mathfrak{a}, \alpha) = \sum_{\pi \equiv \alpha (\text{mod } \mathfrak{a})} |\pi|^2 \leq x, \]
and
\[ R_{\mathfrak{d}}(x) = \Pi((x; 2|1 + i|) - \frac{1}{2} \Pi(x; 2), 1), \]
because $\Pi(x; 2|1 + i|) = \Pi(x; \mathfrak{d}, 1)$ for any $\mathfrak{d}$ odd. Hence, to deduce our bounds for $W_1(x)$, we can use the generalization of Johnson of the Bombieri-Vinogradov theorem, Corollary in page 203 of [J], to imaginary quadratic fields. In particular we have

**Proposition 4** Let $\mathfrak{a}$ run over ideals of $\mathbb{Z}[i]$, and $\alpha$ run over gaussian integers. We have
\[ \sum_{N(\mathfrak{a}) \leq Q} \max_{(\alpha, \mathfrak{a}) = 1} \left| \Pi(x; \mathfrak{a}, \alpha) - \frac{1}{\Phi(\mathfrak{a})} \Pi(x) \right| \ll \frac{x}{(\log x)^A} \quad (5) \]
where $Q = \sqrt{T}/(\log x)^B$ and $\Phi(\mathfrak{a}) = |(\mathbb{Z}[i]/\mathfrak{a})^*|$. Here $A$ is any positive number and $B$ and the implied constant depends only on $A$.

**Proof**: Immediate from the mentioned Corollary in [J].

In our case, $\mathfrak{a} = \delta/(1 + i)^e$, for $e = 1$ or 0 depending on wether $(1 + i) | \mathfrak{d}$ or not, hence
\[ \Phi(\mathfrak{a}) = \prod_{|\pi| \mathfrak{d}} (|\pi|^2 - 1) \prod_{q | \mathfrak{d}} (q^2 - 1), \]
and so
\[ |\mathfrak{A}(x)| = \Pi(x)\hat{g}(\mathfrak{d}) + \hat{r}_{\mathfrak{d}}(\mathfrak{A}(x)), \quad (6) \]
where $\hat{g}(\cdot)$ is the multiplicative function over the ideals in $\mathbb{Z}[i]$ such that
\[ \hat{g}(1 + i) = \frac{1}{2}, \quad \hat{g}(\pi) = \frac{1}{|\pi|^2 - 1}, \quad \hat{g}(q) = \frac{1}{q^2 - 1}. \]
The error terms satisfy
\[ \sum_{N(\mathfrak{d}) \leq \sqrt{T}/(\log x)^B} |\hat{r}_\mathfrak{d} \mathfrak{A}(x)| \ll \frac{x}{(\log x)^A}. \quad (7) \]
Hence, by (3), (4) and (6) we get for \( d = 2^{r}d_1d_2, d_1|P(z), d_2|Q(z) \) as above

\[
|\mathcal{A}_d(x)| = \sum_{k|d} \mu(k)(\Pi(x) \sum_{\mathfrak{d}|k} \hat{\mu}(\mathfrak{d})\hat{g}(\mathfrak{d}) + \sum_{\mathfrak{d}|k} \hat{\mu}(\mathfrak{d})\hat{r}_\vartheta(x)) \\
= \Pi(x) \sum_{k|d} \mu(k)H(k) + \sum_{k|d} \mu(k)\sum_{\mathfrak{d}|k} \hat{\mu}(\mathfrak{d})\hat{r}_\vartheta(x)
\]

where \( H(\cdot) \) is the multiplicative function such that \( H(2) = \frac{1}{2}, H(p) = (1 - \hat{g}(\pi))^2 \) for splitting primes and \( H(q) = 1 - \hat{g}(q) \) for primes inert in \( \mathbb{Z}[i] \). Moreover, by switching the order of summation, we easily get

\[
\sum_{k|d} \mu(k)\sum_{\mathfrak{d}|k} \hat{\mu}(\mathfrak{d})\hat{r}_\vartheta(x) = \sum_{a_2 = d} \hat{\mu}(\mathfrak{d})\hat{r}_\vartheta(x)mu(a_2),
\]

where \( a_2 = 2^{r}a_1b_2 \) whenever \( \mathfrak{d} = (1 + i)^r\alpha_1b_2 \) for \( N(\alpha_1) = a_1, a_1|d_1 \) and \( b_2|d_2 \).

Hence we can write

\[
|\mathcal{A}_d(x)| = \Pi(x)g(d) + r_d(x),
\]

where \( g(\cdot) \) is the multiplicative function such that \( g(l) = 1 - H(l) \) for any prime \( l \), and \( |r_d(x)| \leq \sum_{a_2 = d} |r_\vartheta(x)| \) which gives us, by Proposition 4,

\[
\sum_{x^2d_1d_2 \leq \sqrt{x/(\log x)^A}} |r_d(x)| \ll \frac{x}{(\log x)^A}.
\]

In order to make the level of distribution in the error term as large as possible, we should control the contribution to the sum from moduli \( d \) with \( d_2 \) large. These terms can be easily estimated as follows. First, ignoring that the element in \( \mathcal{A}(x) \) are parametrized by primes, we have

\[
|\mathcal{A}_d(x)| \leq \sum_{u^2 + v^2 \leq x/8} \sum_{d_1u^2 + v^2} \sum_{d_2} 1 \ll \tau(d_1) \frac{x}{d_1d_2^2}.
\]

On the other hand, since \( d \) is squarefree, we have

\[
\Pi(x)g(d) \ll \frac{\tau(d_1)}{\phi(d_1)d_2^2} \frac{x}{\log x},
\]

and so the total contribution to (9) from every \( d \) with \( d_2 \gg (\log x)^{A+1} \) is absorbed by the right hand side. Hence, by changing \( B \) to \( B + A + 1 \), we can write (9) as

\[
\sum_{d \leq \sqrt{x/(\log x)^B}} |r_d(x)| \ll \frac{x}{(\log x)^A}.
\]

Now, a straightforward application of the prime number theorem for the arithmetic progression \( p \equiv 1 \pmod{4} \) allows us to see that the density function \( g(\cdot) \) verifies the linear sieve assumption

\[
\left( \frac{\log z}{\log w} \right) \left( 1 - \frac{L_1}{\log w} \right) \leq \prod_{w \leq p < z} (1 - g(p))^{-1} \leq \left( \frac{\log z}{\log w} \right) \left( 1 + \frac{L_2}{\log w} \right),
\]

(11)
for some constants $L_1, L_2$, and so by (10) and (11) we can apply linear sieve to $A(x)$ with level of distribution $D(x) = \sqrt{x}/(\log x)^B$ to deduce, by the Jurkat-Richert Theorem, that

$$W_1(x) \geq \Pi(x)V(z)f(s)\{1+o(1)\},$$

where $s = \frac{\log D(x)}{\log z}$, $V(z) = \prod_{p \leq z}(1-g(p))$, and $f(s) = 2e^{\gamma \log(x-1)}$ for $2 \leq s \leq 4$, by (5.1) and (5.2) in [I]. In particular, with our selection of $z$ and $D(x)$ above we get $s = 4 - 8B\frac{\log \log x}{\log x}$, $f(s) = \frac{2}{\gamma} \log 3 + o(1)$, and we get the lower bound

$$W_1(x) \geq \left(\frac{1}{2}e^{-\gamma} \log 3 + \varepsilon\right) \Pi(x)V(z),$$

valid for any $\varepsilon > 0$ and for $x$ sufficiently large in terms of $\varepsilon$.

4 Upper bound for $W_2$.

We now proceed to bound $W_2$ from above. Here instead of $A(x)$, the sets to consider in the sieve process would be

$$A_{p_0}(x) = \{a \in A(x) : p_0|a\},$$

for each prime $p_0$ in the interval $(z,y]$. In this case the number of elements in $A_{p_0}$ divisible by $d$ is precisely

$$|A_{dp_0}(x)| = \Pi(x)g(dp_0) + r_{dp_0}(x)$$

for $g(\cdot)$ and $r(\cdot)$ as in (8), and so we can apply the upper-bound linear sieve of Jurkat and Richert, now with level of distribution $D(x)/p_0$, to find

$$\sum_{a \in A_{p_0}(x)} \sum_{d \leq D(x)/p_0} 1 \leq \Pi(x)\sum_{d \leq D(x)/p_0} |r_{dp_0}(x)|,$$

where $s_{p_0} = \log(D(x)/p_0)/\log z$, and $F(s) = 2e^{\gamma}s^{-1}$ for any $1 \leq s \leq 3$ by (3.76) in [I]. In our case, $z < p_0 \leq y$, and so $F(s_{p_0}) = \frac{2}{\gamma} \log(x/p_0) + o(1)$. Hence, summing over all the primes $p_0$ in the interval we get

$$W_2 \leq \Pi(x)V(z) \left\{ \sum_{z < p_0 \leq y} F(s_{p_0})g(p_0) + o(1) \right\},$$

since $\sum_{z < p_0 \leq y} g(p_0) = 2 \sum_{z < p_0 \leq y} \frac{1}{p_0} = O(1)$, and the absolute error term satisfy

$$\sum_{z < p_0 \leq y} \sum_{d \leq D(x)/p_0} |r_{dp_0}(x)| \ll x/(\log x)^A,$
by (10). Partial summation and (11) allow us to obtain

\[ W_2 \leq \Pi(x)V(z) \frac{e^7}{2} \int_{y}^{x} \frac{\log x}{\log(x/t^2)t\log t} dt + o(1). \]

By changing variables \( t = x^u \) we get

\[ W_2 \leq \left( \frac{1}{2} e^7 \log 6 + \varepsilon \right) \Pi(x)V(z), \tag{16} \]

for any \( \varepsilon > 0 \), and \( x \) sufficiently large depending on \( \varepsilon \).

5 Upper bound for \( W_3(x) \).

Finally we have to control \( W_3(x) \) which counts the number of elements \( a \) in \( \mathcal{A}(x) \) such that \( a = p_{1}p_{2}p_{3} \) for splitting primes in certain range. More precisely \( W_3(x) \) counts the total number of solutions to any of the four equations

\[ \pi = 1 + (1+i)^3 \epsilon \pi_1 \pi_2 \pi_3 \text{ with } \epsilon \in \{ \pm 1, \pm i \}, \text{ in primary primes such that} \]

\[ |\pi|^2 \leq x \quad \text{and} \quad z \leq |\pi_3|^2 < y \leq |\pi_2|^2 \leq |\pi_1|^2. \]

For this purpose we will also use a linear sieve and Jurkat-Richert Theorem, not directly, but using a switching device and changing the roles of primes \( \pi \) with the triples \( \pi_1, \pi_2, \pi_3 \). With this in mind let us again note that the condition \(|\pi|^2 \leq x\) can be replaced by \(|\pi_1 \pi_2 \pi_3|^2 \leq x/8\) with a negligible error of \( O(\sqrt[3]{x}) \).

Let us now consider the sequence

\[ B(x) = \{ \Omega(1+\omega) : \omega \in \Omega(x) \}, \]

where

\[ \Omega(x) = \{ \omega = (1+i)^3 \epsilon \pi_1 \pi_2 \pi_3 : \epsilon \in \mathbb{Z}[i]^*, |\omega|^2 \leq x, z \leq |\pi_3|^2 < y < |\pi_2|^2 < |\pi_1|^2 \}. \]

Then, \( W_3(x) \) counts essentially the number of primes in \( B(x) \). In particular

\[ W_3(x) \leq \sum_{\omega \in \Omega(x), \epsilon (1,2P(\sqrt{x})Q(\sqrt{x})) = 1} 1 + O(\sqrt[3]{x}), \]

and the problem is ready to apply sieve theory to the sequence \( B(x) \), in this case with new sieve parameter \( x_0 = \sqrt[3]{x} \). Again we need to estimate \( |B_d(x)| \), the number of elements in \( B(x) \) divisible by \( dP(\sqrt{x})Q(\sqrt{x}) \). Observe that now, if \( 2|d \), then the set \( B_d(x) \) is trivially empty. As before, we will write \( d = d_1 d_2 \) where \( d_1 \mid P(\sqrt{x}) \) and \( d_2 \mid Q(\sqrt{x}) \). Again using Lemma 3 we get

\[ |B_d(x)| = \mu(k) \sum_{k \mid d} S(B(x), k) = \sum_{k \mid d} \mu(k) S(\mathcal{B}(x), k), \]

where

\[ \mathcal{B}(x) = \{ 1 + \omega : \omega \in \Omega(x) \}, \]
and in the same way we will obtain our estimations by aproximations to $S(\mathcal{B}(x), \kappa)$ for any $\kappa |P(\sqrt{z})Q(\sqrt{z})$. For this purpose we note that, similarly as in Section 3,

$$|\mathcal{B}_0(x)| = \sum_{\omega \in \mathcal{I}(x) \cap \{\omega \equiv -1 \pmod{\mathfrak{p}}\}} 1,$$

and so, as in the previous cases, the key point to evaluate $|\mathcal{B}_0(x)|$, and then $|\mathcal{B}_a(x)|$, relies in the existence of an analogous Bombieri-Vinogradov Theorem for the numbers in the set $\Omega(x)$, that we now state. To ease notation we will denote the error term in the approximation as

$$E(x; \alpha, a) = \sum_{\omega \in \mathcal{I}(x) \in \mathcal{A}(\alpha, \mathfrak{a})} 1 - \frac{1}{\Phi(\mathfrak{a})} \sum_{\omega \in \mathcal{I}(x) \in \mathcal{A}(\alpha, \mathfrak{a})} 1.$$

**Proposition 5** Let the notation be as above, and $x > 0$. We have

$$\sum_{\alpha, \mathfrak{a}} \max_{N(\alpha) \leq Q} |E(x; \alpha, a)| \ll \frac{x}{(\log x)^A},$$

with $Q = \sqrt{z}/(\log x)^B$. Here $A$ is any positive number and $B$ and the implied constant depend only on $A$.

**Proof:** First we want to separate the variable $\pi_3$ from the variables $\pi_1, \pi_2$ in the definition of the set $\Omega(x)$. To this end we split $\Omega(x)$ into subsets $\Omega_k(x)$ in which

$$z(1 + \delta)^k \leq |\pi_3| < z(1 + \delta)^{k+1},$$

where $\delta$ is a small number and $0 \leq k \leq K$ with $K = \lfloor \log(y/z)/\log(1 + \delta) \rfloor$. Note that the number of such subsets $\Omega_k(x)$ which cover $\Omega(x)$ is $O(\delta^{-1} \log x)$. In each $\Omega_k(x)$ we replace the condition $|\omega|^2 = 8|\pi_1\pi_2\pi_3|^2 \leq x$ by the condition $8|\pi_1\pi_2|^2 z(1 + \delta)^k \leq x$ and we denote the resulting set by $\Omega'_k(x)$. The above partition and modification cover the set $\Omega(x)$ in a one-to-one fashion, except for the numbers $\omega = (1 + i)^3 \pi_1\pi_2\pi_3$ with $x/(1 + \delta) < 8|\pi_1\pi_2\pi_3|^2 < x(1 + \delta)$ or $y < |\pi_3|^2 \leq (1 + \delta)y$, $8|\pi_1\pi_2|^2 < x$. However this boundary terms contribute to (18) trivially $O(\delta x \log x)^{2006}$, so they can be ignored by choosing $\delta = (\log x)^{-A-2006}$. Therefore, it suffices to show (18) for the restricted sets $\Omega'_k(x)$ separately with $A$ replaced by $2A + 2007$. Put $w = z(1 + \delta)^k$ and let $E_k(x; \alpha, a)$ be the corresponding error term for the set $\Omega'_k(x)$.

By orthogonality of the characters over $(\mathbb{Z}[i]/\mathfrak{a})^*$ we get

$$|E_k(x; \alpha, a)| \leq \frac{1}{\Phi(\mathfrak{a})} \sum_{\chi \not\equiv \chi_0} \left| \sum_{\omega \in \Omega(x)} \chi(\omega) \right| \leq$$

$$\leq \frac{4}{\Phi(\mathfrak{a})} \sum_{\chi \not\equiv \chi_0} \left| \sum_{\omega \leq |\pi_3|^2 \omega(1 + \delta)} \chi(\pi_3) \sum_{y \leq |\pi_2|^2 \leq x/y \omega} \chi(\pi_1\pi_2) \right|.$$
Summing over all ideals of norm up to \( Q \), and splitting into primitive characters we get,
\[
\sum_{N(a) \leq Q} \max_{\alpha} |E(x; \alpha, a)| \ll \sum_{N(a_1) \leq Q} \frac{1}{\Phi(a_1)} \sum_{N(a_2) \leq Q} \frac{1}{\Phi(a_2)} \sum_\chi \left| A_{k,a_1}(\chi)B_{k,a_1}(\chi) \right|. \tag{19}
\]
where
\[
A_{k,a_1}(\chi) = \sum_{(\alpha,a_1)=1} \hat{a}(\alpha)\chi(\alpha), \quad B_{k,a_1}(\chi) = \sum_{(\beta,a_1)=1} \hat{b}(\beta)\chi(\beta),
\]
for
\[
\hat{a}(\alpha) = \begin{cases} 
1 & \text{if } \alpha = \pi_3, w \leq |\pi_3|^2 < w(1 + \delta) \\
0 & \text{otherwise},
\end{cases}
\]
and
\[
\hat{b}(\beta) = \begin{cases} 
1 & \text{if } \beta = \pi_1\pi_2, y \leq |\pi_2| < |\pi_1|, |\beta|^2 < x/8w \\
0 & \text{otherwise}.
\end{cases}
\]
Observe that, in particular, \( \hat{b}(\beta) = 0 \) if \( |\beta|^2 > x/8w \).

The following lemma, generalization of Lemma 17.3 of [I-K], will be crucial in the proof of the proposition. Then, a large sieve inequality in the gaussian domain will allow us to end the proof. Let \( \hat{a}(\alpha) \in \mathbb{C} \), indexed over gaussian integers \( \alpha \) with \( N(\alpha) \leq N \) and such that
\[
\left| \sum_{N(\alpha) \leq N} \hat{a}(\alpha) - \frac{1}{\Phi(\mathfrak{d})} \sum_{N(\alpha) \leq N} \hat{a}(\alpha) \right| \leq ||\hat{a}||N^{1/2}\hat{\tau}(\mathfrak{d}), \tag{20}
\]
for some \( 0 < \Delta < 1 \) and for any ideal \( \mathfrak{d} \in \mathbb{Z}[i] \) and \( \xi \in \mathbb{Z}[i], (\xi,\mathfrak{d}) = 1 \). As usual
\[
||\hat{a}|| = \left( \sum_{N(\alpha) \leq N} |\hat{a}(\alpha)|^2 \right)^{1/2}.
\]

**Lemma 6** Let \( \mathfrak{d} \) be an ideal in \( \mathbb{Z}[i] \), and \( \hat{a}(\alpha) \) a sequence of complex numbers satisfying (20) for any \( N \), and any \( \xi \) modulo \( \mathfrak{d} \), \( (\xi,\mathfrak{d}) = 1 \). Let \( \mathfrak{M} \) be an ideal in \( \mathbb{Z}[i] \) and \( \chi \) a character modulo \( \mathfrak{M} \). Then
\[
\left| \sum_{(\alpha,\mathfrak{d})=1} \hat{a}(\alpha)\chi(\alpha) \right| \leq ||\hat{a}||\Delta^3\hat{\tau}(\mathfrak{d})N(\mathfrak{M})(N \log N)^{1/2},
\]
where \( \hat{\tau}(\mathfrak{d}) \) counts the number of ideal divisors of \( \mathfrak{d} \).
Proof: The proof will go along the lines of Lemma 17.3 in [I-K]. In particular it is easy to see that

\[
\sum_{\alpha, \beta=1} \hat{a}(\alpha)\chi(\alpha) =
\]

\[
= \sum_{N(\epsilon) \leq K} \hat{\mu}(\epsilon) \sum_{\alpha, \beta=1} \hat{a}(\alpha)\chi(\alpha) + \sum_{N(\epsilon) > K} \hat{\mu}(\epsilon) \sum_{\alpha, \beta=1} \hat{a}(\alpha)\chi(\alpha)
\]

\[
= S_1 + S_2,
\]

where \( K \) will be chosen later. By splitting into classes modulo \( \mathfrak{M} \), and applying (20) for each class we get,

\[
S_1 \ll ||\hat{a}||N^{1/2}\Delta^9 \sum_{N(\epsilon) \leq K} \sum_{\alpha \equiv \epsilon (mod \epsilon)} |\hat{\mu}(\epsilon)| \Phi(\mathfrak{M}) \leq K||\hat{a}||N^{1/2}\Delta^9 N(\mathfrak{M})\hat{\tau}(\mathfrak{d}).
\]

For the last inequality we have used that if \( \epsilon \) is a squarefree ideal, then \( \Phi(\mathfrak{M}) \leq \Phi(\epsilon) N(\mathfrak{M}) \) for any ideal \( \mathfrak{M} \) in \( \mathbb{Z}[i] \), and that for any squarefree ideal \( \epsilon \), \( \sum_{\epsilon\mid\epsilon'} \Phi(\epsilon) = N(\epsilon) \). In order to get an upper bound for \( S_2 \) we use twice Cauchy-Schwartz inequality to get

\[
S_2 \leq \hat{\tau}(\mathfrak{d})^{1/2} ||\hat{a}|| \left( \sum_{\epsilon \mid \epsilon'} \sum_{N(\epsilon) \leq N} 1 \right)^{1/2} \leq \hat{\tau}(\mathfrak{d}) ||\hat{a}|| (N \log N)^{1/2} K^{-1/2},
\]

since

\[
\sum_{N(\epsilon) \leq N} 1 \leq \sum_{n \leq N/K} \sum_{\epsilon = n (mod \epsilon)} 1 \leq \sum_{n \leq N/K} \tau(n) \leq \frac{1}{K} N \log N. \quad (21)
\]

The lemma follows by choosing \( K = \Delta^{-6} \).

We now want to use Lemma 6 in (19). For that purpose we split the sum in (19) into two, depending on whether \( N(\alpha_2) \leq R \) or \( N(\alpha_2) > R \). Let us call \( D_1 \) and \( D_2 \) each of these two sums respectively. Since \( ||\hat{a}|| \leq (2w)^{1/2} \), and \( ||\hat{\epsilon}|| \leq (x/(8w))^{1/2} \), we just have to use Lemma 6 to get

\[
D_1 \leq (x \log x) \Delta^3 \sum_{N(\alpha_2) \leq R} \frac{\hat{\tau}(\alpha_1)}{\Phi(\alpha_1)} \sum_{N(\alpha_2) \leq R} N(\alpha_2),
\]

where \( \Delta \) will be chosen so that \( \hat{a} \) verifies (20). We will need two simple lemmas to deal with the sums over ideals \( \alpha_1 \) and \( \alpha_2 \).

**Lemma 7** For any arithmetic function \( f(n) \) such that \( f(nm) \leq f(n)f(m) \) for every pair of integers \( n, m \) we have

\[
\sum_{n \leq x} \tau(n) f(n) \leq \left( \sum_{n \leq x} f(n) \right)^2
\]

for any \( x > 0 \).
Proof: Immediate by switching the order of summation.

Lemma 8 Let \( k \) be a positive integer, and \( x > 0 \).

\[
\sum_{n \leq x} \frac{\tau(n)^k}{\phi(n)} \ll (\log x)^{2k+1}
\]

Proof: Apply the previous lemma to \( \tau(n)^{k-1}/\phi(n) \), induction, and the fact that \( \sum_{n \leq x} \frac{\tau(n)}{\phi(n)} \ll (\log x)^2 \), the last inequality being consequence of the trivial one \( \phi(n) \gg n/\log x \) for any \( n \leq x \).

With this two lemmas on hand it is immediate to get

\[
D_1 \ll x \log x \Delta^3 R^2 (\log R)(\log Q)^8,
\]

(22)

since, (see (21)),

\[
\sum_{N(a) \leq R} N(a_2) \leq R \sum_{N(a_2) \leq R} 1 \ll R^2 \log R,
\]

and, if \( N(c) = n \), then \( \Phi(c) \geq \phi(n) \), and so

\[
\sum_{N(a) \leq Q} \frac{\tau(a)}{\Phi(a)} = \sum_{n \leq Q} \sum_{N(a_2) = n} \frac{\tau(a)}{\Phi(a)} \leq \sum_{n \leq Q} \tau(n) \sum_{N(a_2) = n} \frac{1}{\Phi(a)} \leq \sum_{n \leq Q} \frac{\tau(n)^2}{\phi(n)}.
\]

(23)

Finally we need to estimate \( D_2 \). A direct application of Cauchy-Schwartz gives us

\[
D_2 \leq \sum_{j=0}^{\lfloor \log(Q/R) \rfloor - 1} \sum_{N(a) \leq Q} \frac{1}{\Phi(a)} (A_j B_j)^{1/2}
\]

where

\[
A_j = \sum_{e^j R \leq N(a) \leq e^{j+1} R} \frac{1}{\Phi(a_2)} \sum_{\chi \pmod{a_2}}^{*} A_{k,a_1}(\chi)^2,
\]

\[
B_j = \sum_{e^j R \leq N(a) \leq e^{j+1} R} \frac{1}{\Phi(a_2)} \sum_{\chi \pmod{a_2}}^{*} B_{k,a_1}(\chi)^2.
\]

The estimation of \( A_j, B_j \) is straightforward from the following large sieve inequality in the gaussian domain,

Lemma 9 Let \( \mathfrak{d} \) be an ideal in \( \mathbb{Z}[i] \), \( \hat{a}(\alpha) \) a sequence of complex numbers supported on gaussian integers with \( N(\alpha) \leq N \), and \( Q \geq 1 \). Then

\[
\sum_{N(\alpha) \leq Q} \frac{N(\alpha)}{\Phi(\alpha)} \sum_{\chi \pmod{\alpha} \chi \neq 0}^{*} \left| \sum_{(\alpha,\beta) = 1} \hat{a}(\alpha) \chi(\alpha) \right|^2 \ll (Q^2 + N)||\hat{a}||.
\]
Proof: This is a consequence of (3.1) in page 180 of [H].

We can use the previous lemma to bound $A_j$, $B_j$ and, in this way, deduce that

$$D_2 \ll x^{1/2} \sum_{N(a_1) \leq Q} \frac{1}{\Phi(a_1)} \times$$

$$\times \sum_{j=0}^{\lceil \log(Q/R) \rceil} \frac{1}{\sigma R} (e^{j+1}R + (w)^{1/2}(e^{j+1}R + (x/8w)^{1/2})$$

$$\ll x^{1/2} (\log x)^6 (Q + (x/z)^{1/2} + y^{1/2} + x^{1/2}/R), \quad (24)$$

since

$$\sum_{N(a_1) \leq Q} \frac{1}{\Phi(a_1)} \leq (\log Q)^4,$$

by Lemma (8), as in (23). Now Proposition 5 follows by choosing $\Delta = (\log x)^{-2A - 2013}$, $Q = x^{1/2}/(\log x)^B$ for $B$ given by Proposition 4, and $R = \Delta^{-1}$.

It is straightforward to go from the previous proposition to the estimation

$$\sum_{d^2 \leq \sqrt{x}/(\log x)^A} \|B_d(x)\| = |\Omega(x)|g(d) \ll \frac{x}{(\log x)^A}. \quad (25)$$

Indeed, first note that in our case we have, by (17), $a = \emptyset_1d_2$ with $d_1$, $d_2$ as mentioned above. Then, to get (25) first remove the condition $(\omega, \emptyset_1) = 1$ by noting that the elements in $\Omega(x)$ only have divisors $|\pi|^2 > x^{1/8}$ and so if one of them is fixed, (dividing certain ideal $\emptyset_1$), then we will have trivially less than $x^{7/8}\log x$ elements left, which will be absorbed by the error term. Second, change the summation in terms of the norm by the more convenient in terms of the divisors. The argument to do so is the same as the one done to go from (9) to (10). Observe that, in this case, we have the extra condition $d$ odd since all the elements in $\Omega(x)$ are divisible by 2. Equation (25) allows us to apply linear sieve to the sequence $B_d(x)$ with level of distribution $D(x)$ to obtain, applying one more time Jurkat-Richert Theorem,

$$W_3(x) \leq \prod_{2 < p < \sqrt{x}} (1 - g(p))|\Omega(x)| \{F(1) + o(1)\} = e^\gamma V(z)(\log x)|\{1 + o(1)\},$$

since $F(s) = 2e^\gamma s^{-1}$ as in (14), and $\prod_{2 < p < \sqrt{x}} (1 - g(p)) = \frac{1}{2} V(z)(1 + o(1))$ by (11). To finish the proof we just have to compare $|\Omega(x)|$ with $|\Omega(x)|$ appearing in (13) and (16). By definition we have

$$|\Omega(x)| \leq 4 \sum_{\nu \leq |\pi_1|^2 < |\pi_2|^2 < \sqrt{x}/|\pi_3|} \Pi(x/(8|\pi_3|)|^2)$$

$$\sim \frac{1}{2} \Pi(x) \sum_{\nu \leq |\pi_1|^2 < |\pi_2|^2 < \sqrt{x}/|\pi_3|} \frac{\log x}{\log(x/(|\pi_3|)^2)}.$$
A new application of partial summation, together with a change of variables, as in the deduction of (16), gives

$$|\Omega(x)| \leq \frac{1}{2} \Pi(x) \int_1^x \int_1^{1/2} \frac{1}{1-u-v} \frac{dudv}{uv} = \frac{1}{2} c \Pi(x),$$

for some $c < 0.36308373$. We just have to combine the previous results to get

$$W_3(x) \leq (\frac{1}{2} e^c c + \varepsilon) \Pi(x) V(z).$$

(26)

Theorem 2 follows by plugging (13), (16), and (26) in (2).

References


[I] H. Iwaniec, Sieve methods, notes for a graduate course in Rutgers University, 1996


N. Koblitz, Primality of the number of points on an elliptic curve over a finite field, Pacific J. Math., 131, (1), (1988), 157–165


