International Journal of Number Theory

A NOTE ON SQUARE TOTIENTS

TRISTAN FREIBERG

Department of Mathematics, University of Missouri,
Columbia, MO 65211, USA
freibergt@missouri.edu

CARL POMERANCE

Department of Mathematics, Dartmouth College,
Hanover, NH 03755, USA
carl.pomerance@dartmouth.edu

Received (Day Month Year)
Accepted (Day Month Year)

A well-known conjecture asserts that there are infinitely many primes p for which $p-1$ is a perfect square. We obtain upper and lower bounds of matching order on the number of pairs of distinct primes $p, q \leqslant x$ for which $(p-1)(q-1)$ is a perfect square.

Keywords: Euler's totient function.

Mathematics Subject Classification 2010: 11N37, 11N25, 11N36

1. Introduction

The first of "Landau's problems" on primes is to show that there are infinitely many primes p for which $p - 1 = \square$, that is, a perfect square. Heuristics [5,15] suggest that

$$\#\{p \leqslant x : p - 1 = \square\} \sim \tfrac{1}{2}\mathfrak{S}\int_2^{\sqrt{x}} \frac{\mathrm{d}t}{\log t} \qquad (x \to \infty),$$

where $\mathfrak{S} := \prod_{p>2}\big(1 - (-1/p)/(p-1)\big)$ and $(-1/\cdot)$ is the Legendre symbol. The problem being as unassailable now as it was in 1912 when Landau compiled his famous list, we consider the problem of counting pairs (p, q) of distinct primes for which $(p - 1)(q - 1) = \square$.

Let $\mathbb{P}$ denote the set of all primes and let

$$\mathbf{S} := \{(p, q) \in \mathbb{P} \times \mathbb{P} : p \neq q \text{ and } (p-1)(q-1) = \square\}.$$

For $x \geqslant 2$, let

$$\mathbf{S}(x) := \#\{(p, q) \in \mathbf{S} : p, q \leqslant x\},$$

Theorem 1.1. *There exist absolute constants $c_2 > c_1 > 0$ such that for all $x \geqslant 5$,*

$$c_1 x/\log x < \mathbf{S}(x) < c_2 x/\log x.$$

We remark that the lower bound $\mathbf{S}(x) \gg x/\log x$ gives

$$\mathbf{S}'(x) := \#\{n \leqslant x : n = pq, (p,q) \in \mathbf{S}\} \geqslant \tfrac{1}{2}\mathbf{S}(\sqrt{x}) \gg \sqrt{x}/\log x,$$

improving on the bound $\mathbf{S}'(x) \gg \sqrt{x}/(\log x)^4$ of the first author [10, Theorem 1.2], and independently, [4]. Let ϕ denote Euler's function. Note that for primes p, q we have $\phi(pq) = \square$ if and only if $(p,q) \in \mathbf{S}$. The distribution of integers n with $\phi(n) = \square$ has been considered recently also in [3] and [8, Section 4.8], while the distribution of integers n with n^2 a totient (that is, a value of ϕ) has been considered in [14]. We remark that our proof goes over with trivial modifications to the case of $(p+1)(q+1) = \square$, that is, $\sigma(pq) = \square$, where σ is the sum-of-divisors function. A similar result is to be expected for solutions to $(p+b)(q+b) = \square$ for any fixed nonzero integer b.

In [4,10] solutions to $(p-1)(q-1)(r-1) = m^3$ are also considered, where p, q, r are distinct primes, and more generally $\phi(n) = m^k$, where n is the product of k distinct primes. In [4], the authors show that if the primes in n are bounded by x, there are at least $c_k x/(\log x)^{2k}$ solutions, while in [10], it is shown that there are at least $c_k x/(\log x)^{k+2}$ solutions. Our lower bound construction in the present paper can be extended to give at least $c_k x/(\log x)^{k-1}$ solutions. We do not have a matching upper bound when $k \geqslant 3$.

In addition to notation already introduced, p, q will always denote primes, $\mathbf{1}_{\mathbb{P}}$ denotes the indicator function of $\mathbb{P}$,

$$\pi(x) := \sum_{p \leqslant x} 1, \quad \pi(x; k, b) := \sum_{\substack{p \leqslant x \\ p \equiv b \bmod k}} 1,$$

Λ denotes the von Mangoldt function, μ denotes the Möbius function, $\omega(n)$ denotes the number of distinct prime divisors of n, and $(D/\cdot)$ denotes the Legendre/Kronecker symbol. Note that $A = O(B)$, $A \ll B$ and $B \gg A$ all indicate that $|A| \leqslant c|B|$ for some absolute constant c, $A \asymp B$ means $A \ll B \ll A$, $A = O_\alpha(B)$ and $A \ll_\alpha B$ denote that $|A| \leqslant c(\alpha)|B|$ for some constant c depending on α, and $A \asymp_\alpha B$ denotes that $A \ll_\alpha B \ll_\alpha A$. Also, $A = o(B)$ indicates that $|A| \leqslant c(x)|B|$ for some function $c(x)$ of x that goes to zero as x tends to infinity.

2. Auxiliary lemmas

We will use the following bounds in the proof of Theorem 1.1.

Lemma 2.1. *(i) If $x \geqslant 2$ and $d \geqslant 1$ then*

$$\sum_{n \leqslant x} \frac{1}{\phi(n)} \ll \log x, \quad \sum_{n > x} \frac{1}{\phi(n^2)} \asymp \frac{1}{x}, \quad \text{and} \quad \sum_{\substack{n > x \\ d \mid n^2}} \frac{1}{\phi(n^2)} \ll \frac{d^{1/2}}{\phi(d)x}.$$

(ii) If $n \geqslant 2$ *then*

$$\sum_{m<n} \frac{n^2 - m^2}{\phi(n^2 - m^2)} \ll n.$$

Proof. (i) We have $\sum_{n \leqslant x} 1/n \leqslant 1 + \int_1^x \mathrm{d}t/t = 1 + \log x$, and the first bound follows by using the identity $n/\phi(n) = \sum_{m|n} \mu(m)^2/\phi(m)$ and switching the order of summation. The second bound follows similarly, noting that $\sum_{n>x^2} 1/n^2 \asymp 1/x$ and that $\phi(n^2) = n\phi(n)$. For the third bound, write $d = d_1 d_2^2$, where d_1 is squarefree, and note that $d \mid n^2$ if and only if $d_1 d_2 \mid n$. Thus,

$$\sum_{\substack{n>x \\ d|n^2}} \frac{1}{\phi(n^2)} = \sum_{\substack{n>x \\ d_1 d_2|n}} \frac{1}{n\phi(n)} \leqslant \frac{1}{d_1 d_2 \phi(d_1 d_2)} \sum_{m>x/(d_1 d_2)} \frac{1}{\phi(m^2)}. \tag{2.1}$$

If $d_1 d_2 \leqslant x/2$, this last sum is, by the second part, $O(d_1 d_2/x)$, leading to

$$\sum_{\substack{n>x \\ d|n^2}} \frac{1}{\phi(n^2)} \ll \frac{1}{\phi(d_1 d_2)x} = \frac{d}{\phi(d) d_1 d_2 x} \leqslant \frac{d^{1/2}}{\phi(d)x}.$$

Finally, if $d_1 d_2 > x/2$, the last sum in (2.1) is $O(1)$, leading to

$$\sum_{\substack{n>x \\ d|n^2}} \frac{1}{\phi(n^2)} \ll \frac{1}{d_1 d_2 \phi(d_1 d_2)} \ll \frac{1}{x\phi(d_1 d_2)} \leqslant \frac{d^{1/2}}{\phi(d)x}.$$

(ii) For any positive integer k we have

$$\frac{k}{\phi(k)} = \sum_{\substack{d|k \\ d^2 \leqslant k}} \frac{\mu(d)^2}{\phi(d)} + \sum_{\substack{d|k \\ d^2 > k}} \frac{\mu(d)^2}{\phi(d)} = \sum_{\substack{d|k \\ d^2 \leqslant k}} \frac{\mu(d)^2}{\phi(d)} + O\Big(k^{-1/3}\Big) \ll \sum_{\substack{d|k \\ d^2 \leqslant k}} \frac{\mu(d)^2}{\phi(d)},$$

using the elementary bounds

$$d/\phi(d) \ll \log\log(3d) \quad \text{and} \quad \textstyle\sum_{d|k} \mu(d)^2 = 2^{\omega(k)} = k^{O(1/\log\log k)}.$$

Thus,

$$\sum_{m<n} \frac{n^2 - m^2}{\phi(n^2 - m^2)} \ll \sum_{m<n} \sum_{\substack{d|n^2-m^2 \\ d<n}} \frac{\mu(d)^2}{\phi(d)} = \sum_{d<n} \frac{\mu(d)^2}{\phi(d)} \sum_{\substack{m<n \\ d|n^2-m^2}} 1.$$

If d is squarefree and $d \mid n^2 - m^2$, then $d = d_1 d_2$ for some d_1, d_2 with $n + m \equiv 0 \bmod d_1$ and $n - m \equiv 0 \bmod d_2$. These congruences are satisfied by a unique m modulo $d_1 d_2 = d$, and there are $2^{\omega(d)}$ ways of writing a squarefree integer d as an ordered product of 2 positive integers. Hence

$$\sum_{d<n} \frac{\mu(d)^2}{\phi(d)} \sum_{\substack{m<n \\ d|n^2-m^2}} 1 = \sum_{d<n} \frac{\mu(d)^2}{\phi(d)} \sum_{d_1 d_2 = d} \sum_{\substack{m<n \\ d_1|n+m \\ d_2|n-m}} 1 \ll n \sum_{d<n} \frac{\mu(d)^2 2^{\omega(d)}}{d\phi(d)} \ll n.$$

$\square$

We will need uniform bounds for $\pi(x;k,b)$ for k up to a small power of x. The following form of the Brun–Titchmarsh inequality is a consequence of a sharp form of the large sieve inequality due to Montgomery and Vaughan [13].

Lemma 2.2. *If $1 \leqslant k < x$ and $(b,k) = 1$ then*

$$\pi(x;k,b) < \frac{2x}{\phi(k)\log(x/k)}.$$

Proof. See [13, Theorem 2]. □

We do not have a matching lower bound for all k up to a power of x because of putative Siegel zeros, however these only affect a very few moduli k that are multiples of certain "exceptional" moduli.

Lemma 2.3. *For any given $\epsilon, \delta > 0$, there exist numbers $\eta_{\epsilon,\delta} > 0$, $x_{\epsilon,\delta}$, $D_{\epsilon,\delta}$ such that whenever $x \geqslant x_{\epsilon,\delta}$, there is a set $\mathcal{D}_{\epsilon,\delta}(x)$, of at most $D_{\epsilon,\delta}$ integers, for which*

$$\left|\pi(x;k,b) - \frac{x}{\phi(k)\log x}\right| \leqslant \frac{\epsilon x}{\phi(k)\log x}$$

whenever k is not a multiple of any element of $\mathcal{D}_{\epsilon,\delta}(x)$, k is in the range

$$1 \leqslant k \leqslant x^{-\delta+5/12},$$

and $(b,k) = 1$. Furthermore, every integer in $\mathcal{D}_{\epsilon,\delta}(x)$ exceeds $\log x$, and all, but at most one, exceed $x^{\eta_{\epsilon,\delta}}$.

Proof. See [1, Theorem 2.1]. □

In fact we will need to count primes $p \equiv b \bmod k$ for which the quotient $(p-b)/k$ is squarefree. We apply an inclusion-exclusion argument to Lemma 2.3.

Lemma 2.4. *There exist absolute constants $\eta > 0$, x_0, D such that whenever $x \geqslant x_0$, there is a set $\mathcal{D}(x)$, of at most D integers, for which*

$$\sum_{a \leqslant x/k} \mu(a)^2 \mathbf{1}_{\mathbb{P}}(ak+b) > \frac{x}{100\phi(k)\log x}$$

whenever $36k$ is not a multiple of any element of $\mathcal{D}(x)$, k is in the range $1 \leqslant k \leqslant x^{1/3}$, and $(b,k) = 1$ with $1 \leqslant b < k$. Furthermore, every integer in $\mathcal{D}(x)$ exceeds $\log x$, and all, but at most one, exceed x^{η}.

Proof. Let $1 \leqslant b < k \leqslant x^{1/3}$ with $(b,k) = 1$. Using $\mu(a)^2 \geqslant 1 - \sum_{p^2|a} 1$ and switching the order of summation, we obtain

$$\begin{aligned}\sum_{a \leqslant x/k} \mu(a)^2 \mathbf{1}_{\mathbb{P}}(ak+b) &\geqslant \sum_{a \leqslant x/k} \mathbf{1}_{\mathbb{P}}(ak+b) - \sum_{p \leqslant \sqrt{x/k}} \sum_{c \leqslant x/(p^2k)} \mathbf{1}_{\mathbb{P}}(cp^2k+b) \\ &\geqslant \pi(x;k,b) - \sum_{p \leqslant \sqrt{x/k}} \pi(x;p^2k,b) - \sqrt{x/k}.\end{aligned}$$

Let $1 \leqslant y < z < \sqrt{x/k}$. Trivially, we have

$$\sum_{z<p\leqslant\sqrt{x/k}} \pi(x;p^2k,b) \leqslant \sum_{p>z} \frac{x}{p^2k} \ll \frac{x}{kz\log z}.$$

Here we have used the bound $\sum_{p>z} 1/p^2 \ll 1/(z\log z)$, which follows from the bound $\pi(x) \ll x/\log x$ by partial summation. By Lemma 2.2 we have

$$\sum_{y<p\leqslant z} \pi(x;p^2k,b) < \frac{2x}{\log(x/(z^2k))}\sum_{p>y}\frac{1}{\phi(p^2k)} \leqslant \frac{2x}{\phi(k)\log(x/(z^2k))}\sum_{p>y}\frac{1}{p(p-1)},$$

using $\phi(p^2k) \geqslant \phi(p^2)\phi(k)$.

We set $y=3$ and $z=\log x$ so that $\log(x/(z^2k)) \sim \log(x/k) \geqslant \frac{2}{3}\log x$. We verify that $\sum_{p>3} 1/(p(p-1)) < 0.1065$. Combining everything gives

$$\sum_{a\leqslant x/k} \mu(a)^2 \mathbf{1}_{\mathbb{P}}(ak+b) > \pi(x;k,b) - \pi(x;4k,b) - \pi(x;9k,b) - \frac{0.32x}{\phi(k)\log x}$$

for all sufficiently large x. We complete the proof by applying Lemma 2.3 with $\epsilon = 1/1000$ and $\delta = 1/12$, noting that $1-1/2-1/6-3\epsilon-0.32 > 1/100$. $\square$

We remark that with more work, a version of Lemma 2.4 can be proved as an equality, with the factor $1/100$ replaced with $c_k + o(1)$ (as $x\to\infty$), where c_k is Artin's constant $\prod_p(1-1/(p(p-1))$ times $\prod_{p|k}(1-1/(p^3-p^2-p))$.

Lemma 2.5. *Fix $\delta \in (0,1]$ and let $x \geqslant 3$. There is a set $\mathcal{E}_\delta(x)$ of quadratic, primitive characters, all of conductor less than x, satisfying $\#\mathcal{E}_\delta(x) \ll_\delta x^\delta$ and such that the following holds. If χ is a real, primitive character of conductor $d \leqslant x$ and $\chi \notin \mathcal{E}_\delta(x)$, then*

$$\prod_{y<p\leqslant z}\left(1-\frac{\chi(p)}{p}\right) \asymp_\delta 1$$

uniformly for $z > y \geqslant \log x$.

Proof. See [6, Lemma 3.3]. The authors of [6] state that the proof of their lemma borrows from [11, Proposition 2.2], and the authors of [11] state that their proposition is essentially due to Elliott [9]. (The lemma, as stated here, is quoted from [14, Lemma 7], and is equivalent to [6, Lemma 3.3].) $\square$

Lemma 2.6. *If $x \geqslant 2$ then*

$$\sum_{a\leqslant x}\frac{a\mu(a)^2}{\phi(a)^2}\prod_{2<p\leqslant\sqrt{x}}\left(1-\frac{(-a/p)}{p}\right)^2 \ll \log x.$$

Proof. First, we note that for $y \geqslant 1$ we have the elementary bound

$$\sum_{a>y}\frac{a^2}{\phi(a)^4} \ll \frac{1}{y}. \tag{2.2}$$

To see this, let h be the multiplicative function satisfying $a^4/\phi(a)^4 = \sum_{m|a} h(m)$, so that

$$h(m) = \mu(m)^2 \prod_{p|a} \left(\frac{p^4}{p^4-1} - 1\right).$$

Then

$$\sum_{a>y} \frac{a^2}{\phi(a)^4} = \sum_{a>y} \frac{1}{a^2}\frac{a^4}{\phi(a)^4} = \int_y^\infty \frac{2}{t^3} \sum_{y<a\leqslant t} \frac{a^4}{\phi(a)^4}\,\mathrm{d}t$$
$$\leqslant \int_y^\infty \frac{2}{t^2} \sum_{m\leqslant t} \frac{h(m)}{m}\,\mathrm{d}t < \frac{2}{y} \sum_{m\geqslant 1} \frac{h(m)}{m}.$$

This last sum has a convergent Euler product, so (2.2) is established.

For a positive squarefree integer a, let χ_a be the Dirichlet character that sends an odd prime p to $(-a/p)$, and such that $\chi_a(2) = 1$ or 0 depending on whether $a \equiv 3 \bmod 4$ or not, respectively. The character χ_a is primitive and has conductor a if $a \equiv 3 \bmod 4$ and $4a$ otherwise.

The product in the lemma (without being squared) resembles $L(1,\chi_a)^{-1}$, in fact,

$$L(1,\chi_a)^{-1} = \prod_p \left(1 - \frac{(-a/p)}{p}\right).$$

Our first goal is to show that we uniformly have

$$L(1,\chi_a) \prod_{2<p\leqslant\sqrt{x}} \left(1 - \frac{(-a/p)}{p}\right) \ll 1 \tag{2.3}$$

for all small a and most other values of $a \leqslant x$. Suppose that $a \leqslant (\log x)^4$. Considering the $\phi(4a)$ residue classes $r \bmod 4a$ that are coprime to $4a$, we see (since the conductor of χ_a divides $4a$) that $(-a/p) = 1$ or -1 depending on which class p lies in, with $\frac{1}{2}\phi(4a)$ classes giving 1 and $\frac{1}{2}\phi(4a)$ classes giving -1. It follows from the Siegel–Walfisz theorem [7, §22 (4)] that

$$\sum_{p>\sqrt{x}} \frac{(-a/p)}{p} = \int_{\sqrt{x}}^\infty \frac{1}{t^2} \sum_{\sqrt{x}<p\leqslant t} (-a/p)\,\mathrm{d}t \ll \phi(4a) \int_{\sqrt{x}}^\infty \frac{1}{t(\log t)^5}\,\mathrm{d}t \ll 1.$$

Exponentiating, we get (2.3).

Now suppose that $a > (\log x)^4$. We break the interval $((\log x)^4, x]$ into dyadic intervals of the form $I_j := [2^j, 2^{j+1})$, where the first and last intervals may overshoot a bit. Using Lemma 2.5 with $\delta = \frac{1}{4}$, $y = \sqrt{x}$, and letting $z \to \infty$, we have (2.3) for all $a \in I_j$ except for possibly $O\big(2^{j/4}\big)$ values of a. Using the trivial estimate

$$\prod_{2<p\leqslant\sqrt{x}} \left(1 - \frac{(-a/p)}{p}\right)^2 \ll (\log x)^2$$

and $a/\phi(a)^2 \ll (\log\log a)^2 a^{-1}$, the contribution of these exceptional values of $a \in I_j$ to the sum in the lemma is

$$\ll 2^{j/4}(\log j)^2 2^{-j}(\log x)^2,$$

which when summed over integers j being considered gives $O\big((\log\log x)^2/\log x\big)$. Thus, we may ignore these exceptional values of a, so assuming that (2.3) always holds.

By the Cauchy–Schwarz inequality we have

$$\sum_{a\in I_j} \frac{a\mu(a)^2}{\phi(a)^2} L(1,\chi_a)^{-2} \leqslant \Bigg(\sum_{a\in I_j} \frac{a^2}{\phi(a)^4}\Bigg)^{1/2} \Bigg(\sum_{a\in I_j} \mu(a)^2 L(1,\chi_a)^{-4}\Bigg)^{1/2}.$$

Now the first sum is $O\big(2^{-j}\big)$ by (2.2), and the second sum is $O\big(2^{j}\big)$ by [11, Theorem 2] (with $z = -4$) and the subsequent comment about Siegel's theorem. Thus, the contribution from $a \in I_j$ to the sum in the lemma is $O(1)$, and since there are $O(\log x)$ choices for j, the lemma is proved. $\square$

We remark that [2, Section 10] has a similar calculation as in Lemma 2.6.

3. Proof of Theorem 1.1

Our proof begins with the observation that every positive integer has a unique representation of the form an^2, where a and n are positive integers with a squarefree. Thus, $(p-1)(q-1) = \square$ if and only if $p = am^2+1$ and $q = an^2+1$ for some squarefree a. It follows that for all $x \geqslant 0$,

$$\mathbf{S}(x+1) = \sum_{a\leqslant x} \mu(a)^2 \sum_{\substack{m,\, n\leqslant\sqrt{x/a}\\ m\neq n}} \mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1). \tag{3.1}$$

3.1. *The lower bound*

Let $x \geqslant 4$ and consider a dyadic interval

$$I_y := [y/2, y) \subset [1, x^{1/6}].$$

Also let

$$N_{I_y}(a) := \sum_{n\in I_y} \mathbf{1}_{\mathbb{P}}(an^2+1). \tag{3.2}$$

Letting $\mathscr{I}$ denote a collection of disjoint dyadic intervals I_y, we deduce from (3.1) that

$$\mathbf{S}(x+1) \geqslant \sum_{I_y\in\mathscr{I}} \sum_{a\leqslant x/y^2} \mu(a)^2 (N_{I_y}(a)^2 - N_{I_y}(a)). \tag{3.3}$$

By the Cauchy–Schwarz inequality, for every $I_y \in \mathscr{I}$ we have

$$\Bigg(\sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y}(a)\Bigg)^2 \leqslant \frac{x}{y^2} \sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y}(a)^2. \tag{3.4}$$

Lemma 3.1. *Given an interval $I_y = [y/2, y)$ and an integer a, let $N_{I_y}(a)$ be as in* (3.2). (*i*) *Uniformly for $2 \leqslant y < \sqrt{x}$, we have*

$$\sum_{a \leqslant x/y^2} N_{I_y}(a) \ll \frac{x}{y \log(x/y^2)}.$$

(*ii*) *Uniformly for $2 \leqslant y \leqslant x^{1/6}$, we have*

$$\sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y}(a) \gg \frac{x}{y \log x}.$$

Proof. (i) We change the order of summation and apply Lemma 2.2:

$$\sum_{a \leqslant x/y^2} N_{I_y}(a) = \sum_{n \in I_y} \sum_{a \leqslant x/y^2} \mathbf{1}_{\mathbb{P}}(an^2+1) \ll \sum_{n \in I_y} \pi(x; n^2, 1) \ll \sum_{n \in I_y} \frac{x}{\phi(n^2)\log(x/n^2)}.$$

We have $\sum_{n \in I_y} 1/\phi(n^2) \ll 1/y$ by the second bound in Lemma 2.1 (i).

(ii) Let $2 \leqslant y \leqslant x^{1/6}$ and let I_y' be the subset of those $n \in I_y$ for which

$$\sum_{a \leqslant x/n^2} \mu(a)^2 \mathbf{1}_{\mathbb{P}}(an^2+1) > \frac{x}{100\phi(n^2)\log x}.$$

Letting $N_{I_y'}(a) := \sum_{n \in I_y'} \mathbf{1}_{\mathbb{P}}(an^2+1)$ we see, after switching the order of summation, that

$$\sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y}(a) \geqslant \sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y'}(a) \geqslant \sum_{n \in I_y'} \sum_{a \leqslant x/n^2} \mu(a)^2 \mathbf{1}_{\mathbb{P}}(an^2+1),$$

and hence

$$\sum_{a \leqslant x/y^2} \mu(a)^2 N_{I_y}(a) > \frac{x}{100\log x} \sum_{n \in I_y'} \frac{1}{\phi(n^2)}.$$

We claim that

$$\sum_{n \in I_y'} \frac{1}{\phi(n^2)} \gg \frac{1}{y}, \tag{3.5}$$

whence the result. The claim follows from the second bound in Lemma 2.1 (i) if $I_y' = I_y$, so let us assume that $I_y' \subsetneq I_y$.

If $n \in I_y \setminus I_y'$ then $n^2 \leqslant x^{1/3}$, and so if x is sufficiently large (as we assume), $36n^2$ is a multiple of an element of the "exceptional set" $\mathcal{D}(x)$ of Lemma 2.4. Hence, by the third bound in Lemma 2.1 (i),

$$\sum_{n\in I_y\setminus I'_y}\frac{1}{\phi(n^2)}\leqslant\sum_{d\in\mathcal{D}(x)}\sum_{\substack{n\in I_y\\ d|36n^2}}\frac{1}{\phi(n^2)}\ll\sum_{d\in\mathcal{D}(x)}\sum_{\substack{n\in I_y\\ d|(6n)^2}}\frac{1}{\phi((6n)^2)}$$
$$\leqslant\sum_{d\in\mathcal{D}(x)}\sum_{\substack{m\geqslant 3y\\ d|m}}\frac{1}{\phi(m^2)}\ll\frac{1}{y}\sum_{d\in\mathcal{D}(x)}\frac{d^{1/2}}{\phi(d)}\ll\frac{\log\log x}{y(\log x)^{1/2}},$$

where the last bound holds because, by Lemma 2.4, there are at most D elements in $\mathcal{D}(x)$, and all elements in $\mathcal{D}(x)$ are greater than $\log x$. Since our estimate is $o(1/y)$ as $x\to\infty$, we have (3.5), and so the lemma. □

Deduction of the lower bound. Combining (3.4) with Lemma 3.1 (i) and (ii), we see that if $I_y=[y/2,y)$, then, uniformly for $(\log x)^2\leqslant y\leqslant x^{1/6}$,

$$\sum_{a\leqslant x/y^2}\mu(a)^2(N_{I_y}(a)^2-N_{I_y}(a))\geqslant\frac{y^2}{x}\bigg(\sum_{a\leqslant x/y^2}\mu(a)^2N_{I_y}(a)\bigg)^2-\sum_{a\leqslant x/y^2}N_{I_y}(a)$$
$$\gg\frac{x}{(\log x)^2}.$$

Letting $\mathscr{I}=\{[2^{j-1},2^j):(\log x)^2\leqslant 2^j\leqslant x^{1/6}\}$ and applying (3.3), we conclude that

$$\mathbf{S}(x)\gg\sum_{I_y\in\mathscr{I}}\frac{x}{(\log x)^2}\gg\frac{x}{\log x}.$$

□

3.2. *The upper bound*

By (3.1) we have $\mathbf{S}(x+1)=2\mathbf{S}_1(x)+2\mathbf{S}_2(x)$, where

$$\mathbf{S}_1(x):=\sum_{a\leqslant x^{2/3}}\mu(a)^2\sum_{n\leqslant\sqrt{x/a}}\sum_{m<n}\mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1)$$
$$\leqslant\sum_{a\leqslant x^{2/3}}\mu(a)^2\bigg(\sum_{n\leqslant\sqrt{x/a}}\mathbf{1}_{\mathbb{P}}(an^2+1)\bigg)^2\tag{3.6}$$

and

$$\mathbf{S}_2(x):=\sum_{x^{2/3}<a\leqslant x}\mu(a)^2\sum_{n\leqslant\sqrt{x/a}}\sum_{m<n}\mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1)$$
$$\leqslant\sum_{n<x^{1/6}}\sum_{m<n}\sum_{a\leqslant x/n^2}\mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1).\tag{3.7}$$

Lemma 3.2. (*i*) *Uniformly for $x\geqslant 2$ and $1\leqslant a\leqslant x^{2/3}$, we have*

$$\sum_{n\leqslant\sqrt{x/a}}\mathbf{1}_{\mathbb{P}}(an^2+1)\ll\frac{\sqrt{x/a}}{\log x}\frac{a}{\phi(a)}\prod_{2<p\leqslant\sqrt{x}}\left(1-\frac{(-a/p)}{p}\right).$$

(ii) Uniformly for $1 \leqslant m < n < x^{1/3}$*, we have*

$$\sum_{a \leqslant x/n^2} \mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1) \ll \frac{x}{(n\log x)^2} \cdot \frac{(m,n)}{\phi((m,n))} \cdot \frac{n^2-m^2}{\phi(n^2-m^2)}.$$

Proof. (i) Given $x \geqslant 2$ and $1 \leqslant a \leqslant x^{2/3}$, let

$$\rho_a(p) := \#\{b \bmod p : ab^2+1 \equiv 0 \bmod p\}.$$

A routine application of Brun's sieve [12, Theorem 2.2] gives

$$\sum_{n \leqslant \sqrt{x/a}} \mathbf{1}_{\mathbb{P}}(an^2+1) \ll \sqrt{x/a} \prod_{p \leqslant \sqrt{x}} \left(1 - \frac{\rho_a(p)}{p}\right).$$

Since $1-\rho_a(p)/p = (1-1/p)(1-(\rho_a(p)-1)/(p-1))$, Mertens' theorem gives

$$\prod_{p \leqslant \sqrt{x}} \left(1 - \frac{\rho_a(p)}{p}\right) \ll \frac{1}{\log x} \prod_{2 < p \leqslant \sqrt{x}} \left(1 - \frac{\rho_a(p)-1}{p-1}\right).$$

Now, $\rho_a(p)-1 = (-a/p)$ for odd $p \nmid a$, and $\rho_a(p) = 0$ for $p \mid a$, hence

$$\prod_{2 < p \leqslant \sqrt{x}} \left(1 - \frac{\rho_a(p)-1}{p-1}\right) \leqslant \frac{a}{\phi(a)} \prod_{2 < p \leqslant \sqrt{x}} \left(1 - \frac{(-a/p)}{p-1}\right),$$

which proves the inequality in the lemma with $p-1$ in the denominator instead of p. But $1-(-a/p)/(p-1) = (1-(-a/p)/p)\big(1+O\big(1/p^2\big)\big)$ so the bound in the lemma holds.

(ii) Given $1 \leqslant m < n < x^{1/3}$, let

$$\rho_{m,n}(p) := \#\{b \bmod p : (bm^2+1)(bn^2+1) \equiv 0 \bmod p\}.$$

Again by Brun's sieve [12, Theorem 2.2],

$$\sum_{a \leqslant x/n^2} \mathbf{1}_{\mathbb{P}}(am^2+1)\mathbf{1}_{\mathbb{P}}(an^2+1) \ll \frac{x}{n^2} \prod_{p \leqslant \sqrt{x}} \left(1 - \frac{\rho_{m,n}(p)}{p}\right).$$

By Mertens' theorem we have

$$\begin{aligned}\prod_{p \leqslant \sqrt{x}} \left(1 - \frac{\rho_{m,n}(p)}{p}\right) &= \prod_{p \leqslant \sqrt{x}} \left(1 + \frac{p(2-\rho_{m,n}(p))-1}{(p-1)^2}\right)\left(\frac{p-1}{p}\right)^2 \\ &\ll \frac{1}{(\log x)^2} \prod_{p \leqslant \sqrt{x}} \left(1 + \frac{p(2-\rho_{m,n}(p))-1}{(p-1)^2}\right).\end{aligned}$$

Now, for any prime p we have

$$\rho_{m,n}(p) = \begin{cases} 2 & \text{if } p \nmid mn(m^2-n^2), \\ 1 & \text{if } p \mid mn(m^2-n^2) \text{ and } p \nmid (m,n), \\ 0 & \text{if } p \mid (m,n), \end{cases}$$

hence

$$\prod_{p\leqslant\sqrt{x}}\left(1+\frac{p(2-\rho_{m,n}(p))-1}{(p-1)^2}\right)\leqslant\prod_{p\mid(m,n)}\left(\frac{p}{p-1}\right)^2\prod_{\substack{p\mid m^2-n^2\\ p\nmid(m,n)}}\frac{p}{p-1}$$
$$=\prod_{p\mid(m,n)}\frac{p}{p-1}\prod_{p\mid(m^2-n^2)}\frac{p}{p-1}.$$

Combining gives the result. □

Deduction of the upper bound. By (3.6), Lemma 3.2 (i) and Lemma 2.6, we have

$$\mathbf{S}_1(x)\ll\frac{x}{(\log x)^2}\sum_{a\leqslant x^{2/3}}\frac{a\mu(a)^2}{\phi(a)^2}\prod_{2<p\leqslant\sqrt{x}}\left(1-\frac{(-a/p)}{p}\right)^2\ll\frac{x}{\log x}.$$

By (3.7) and Lemma 3.2 (ii) we have

$$\mathbf{S}_2(x)\ll\frac{x}{(\log x)^2}\sum_{n<x^{1/6}}\frac{1}{n^2}\sum_{m<n}\frac{(m,n)}{\phi((m,n))}\cdot\frac{n^2-m^2}{\phi(n^2-m^2)}.$$

To bound the double sum, we write $g=(m,n)$, $m=gm_1$, $n=gn_1$ and change the order of summation to obtain

$$\sum_{g\leqslant x^{1/6}}\frac{1}{g^2}\sum_{n_1\leqslant x^{1/6}/g}\frac{1}{n_1^2}\sum_{\substack{m_1<n_1\\(m_1,n_1)=1}}\frac{g}{\phi(g)}\cdot\frac{g^2(n_1^2-m_1^2)}{\phi(g^2(n_1^2-m_1^2))}$$
$$\leqslant\sum_{g\leqslant x^{1/6}}\frac{1}{\phi(g)^2}\sum_{n_1\leqslant x^{1/6}/g}\frac{1}{n_1^2}\sum_{\substack{m_1<n_1\\(m_1,n_1)=1}}\frac{n_1^2-m_1^2}{\phi(n_1^2-m_1^2)}.$$

This is equal to $O\Big(\sum_{n_1\leqslant x}1/n_1\Big)=O(\log x)$ by Lemma 2.1 (ii). Recalling that $\mathbf{S}(x)=2\mathbf{S}_1(x)+2\mathbf{S}_2(x)$ and combining gives

$$\mathbf{S}(x)\ll\mathbf{S}_1(x)+\mathbf{S}_2(x)\ll\frac{x}{\log x}.$$

This completes the proof of the theorem. □

Acknowledgments

The work on this paper commenced when the second author was visiting the first at KTH, Stockholm. Both authors gratefully acknowledge the hospitality of Pär Kurlberg.

References

[1] W. R. Alford, A. Granville and C. Pomerance, There are infinitely many Carmichael numbers, *Ann. of Math. (2)* **139**(3) (1994) 703–722.

[2] S. Baier and L. Zhao, On primes in quadratic progressions, *Int. J. Number Theory* **5**(6) (2009) 1017–1035.

[3] W. D. Banks, J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, Multiplicative structure of values of the Euler function, in *High Primes and Misdemeanours: Lectures in Honour of the Sixtieth Birthday of Hugh Cowie Williams*, eds. A. Van der Poorten and A. Stein, Fields Inst. Comm., Vol 41 (American Mathematical Society, 2004), pp. 29–47.

[4] W. D. Banks and F. Luca, Power totients with almost primes, *Integers* **11**(3) (2011), 307–313.

[5] P. T. Bateman and R. A. Horn, A heuristic asymptotic formula concerning the distribution of prime numbers, *Math. Comp.* **16**(79) (1962) 363–367.

[6] V. Chandee, C. David, D. Koukoulopoulos and E. Smith, Group structures of elliptic curves over finite fields, *Int. Math. Res. Not.* To appear.

[7] H. Davenport, *Multiplicative number theory*, 3rd edn., Graduate Texts in Mathematics No. 74 (Springer–Verlag, 2000). Revised and with a preface by H. L. Montgomery.

[8] J.-M. De Koninck and F. Luca, *Analytic number theory: exploring the anatomy of integers* (American Mathematical Society, 2012).

[9] P. D. T. A. Elliott, On the size of $L(1, \chi)$, *J. Reine Angew. Math.* **236** (1969) 26–36.

[10] T. Freiberg, Products of shifted primes simultaneously taking perfect power values, *J. Aust. Math. Soc.* **92**(2) (2012) 145–154.

[11] A. Granville and K. Soundararajan, The distribution of values of $L(1, \chi_d)$, *Geom. Funct. Anal. (13)* **13**(5) (2003) 992–1028.

[12] H. Halberstam and H. E. Richert, *Sieve methods*, London Mathematical Society Monographs, No. 4. (Academic Press, 1974).

[13] H. L. Montgomery and R. C. Vaughan, The large sieve, *Mathematika* **20**(2) (1973) 119–134.

[14] P. Pollack and C. Pomerance, Square values of Euler's function, *Bull. Lond. Math. Soc.* **46**(2) (2014) 403–414.

[15] D. Shanks, On the conjecture of Hardy & Littlewood concerning the number of primes of the form $n^2 + a$, *Math. Comp.* **14**(72) (1960) 320–332.