ACTA ARITHMETICA
Online First version

Two problems on the greatest prime factor of $n^2 + 1$

by

Glyn Harman (Egham)

Abstract. Let $P^+(m)$ denote the greatest prime factor of the positive integer m. In [Arch. Math. (Basel) 90 (2008), 239–245] we improved work of Dartyge [Acta Math. Hungar. 72 (1996), 1–34] to show that

$$|\{n \le x : P^+(n^2+1) < x^\alpha\}| \gg x$$

for $\alpha > 4/5$. In this note we show how the recent work of de la Bretèche and Drappeau [J. Eur. Math. Soc. 22 (2020), 1577–1624] (which uses the improved bound for the smallest eigenvalue in the Ramanujan–Selberg conjecture given by Kim [J. Amer. Math. Soc. 16 (2003), 139–183]) along with a change of argument can be used to reduce the exponent to 0.567. We also show how recent work of Merikoski [J. Eur. Math. Soc. 25 (2023), 1253–1284] on large values of $P^+(n^2+1)$ can improve work by Everest and the author [London Math. Soc. Lecture Note Ser. 352, Cambridge Univ. Press, 2008, 142–154] on primitive divisors of the sequence n^2+1.

1. Introduction. On page 23 of [12] the following conjecture is asserted.

There are infinitely many primes $n^2 + 1$. More generally, if a, b, c are integers without common divisor, a is positive, $a + b$ and c are not both even, and $b^2 - 4ac$ is not a perfect square, then there are infinitely many primes $an^2 + b + c$.

Indeed, there is a more general conjecture on irreducible polynomials without fixed prime divisors, and this has been put into a quantitative form [2, 11]. In the same way, conjectures have been made on "smooth" values of polynomials. For example, it is reasonable, given $-D$ not an integer squared, to suppose that, given $\epsilon > 0$, one should have $P^+(n^2 + D) < n^\epsilon$ infinitely often, where $P^+(m)$ denotes the greatest prime factor of the positive integer m. Indeed, this has been proved, in a slightly stronger form with an explicit $\epsilon(n, D) \to 0$, by Schinzel [20, Theorem 13]. However, Schinzel's method does not give the expected formula for the number of such values of $n^2 + D$.

2020 *Mathematics Subject Classification*: Primary 11N32.
Key words and phrases: prime, quadratic polynomial.
Received 10 July 2023; revised 15 December 2023.
Published online 21 February 2024.

In fact, given $\alpha > 0$ (not necessarily "small") it cannot even provide a lower bound of the correct order of magnitude for

$$|\{n \le x : P^+(n^2 + D) < x^\alpha\}| = \Psi_D^*(x, \alpha), \quad \text{say},$$

since the values of n are stated explicitly in a form $\ge 2^m$ for a certain sequence of values m [20, p. 230].

There is more recent work which covers general quadratic polynomials [4], but still cannot provide the correct order lower bound. An asymptotic formula has been given by Martin [18], but only dependent on a very strong unproved hypothesis (a uniform quantitative version of "Hypothesis H" of Schinzel and Sierpiński). In 1996 Dartyge [5] proved that $\Psi_1^*(x, \alpha) \gg x$ for $\alpha > \frac{149}{179} \approx 0.8324$. The method drew on the techniques used in [14, 7] for proving that $P^+(n^2 + 1) > n^\gamma$ with $\gamma > 1$, combined with methods of Balog [1] and Friedlander [10] for obtaining the correct order of magnitude of smooth numbers in certain sequences.

In [13] we dispensed with Balog's method and thereby reduced the lower bound for α to $\frac{4}{5}$, and the first purpose of the present paper is to give the following more significant improvement. The value 0.567 which occurs in our main result arises as

$$\frac{356}{381} e^{-1/2} + \epsilon$$

for any $\epsilon > 0$. Here, and throughout the paper, we reserve the letter e for the base of natural logarithms. There are no serious mathematical problems in replacing $n^2 + 1$ with $n^2 + D$ (where $-D$ is not an integer squared) in any of our results. We have restricted ourselves to $n^2 + 1$ for brevity and clarity. All the main lemmas have exact analogues for $n^2 + D$ which can be derived using Hooley's work [14, 15] and noting that [3] deals with the more general case. Henceforth we shall therefore suppress the subscript 1 on $\Psi_1^*(x, \alpha)$.

THEOREM 1.1. *For* $\alpha \ge 0.567$ *we have*

$$\Psi^*(x, \alpha) \gg x. \tag{1}$$

If we fed an improved result from [3] into the method of [13] we would only get an exponent of $82/107 + \epsilon \approx 0.766$. However, we shall use a different result from [3] and combine that with the Balog–Friedlander approach to get a much better improvement in the exponent (though paradoxically this makes our method resemble Dartyge's approach [5]). We remark that Merikoski [19] has combined the work of de la Brèteche and Drappeau with other ideas to show that infinitely often $P^+(n^2 + 1) > n^{1.279}$. The methods used to prove these types of results have implications for the work given in [8, 9] on primitive divisors of quadratic polynomials, and we shall briefly describe one such result in our final section.

2. Outline of the method. Write

$$\phi = e^{-1/2}, \quad \beta = \frac{356}{381}, \quad \alpha = \beta\phi + \epsilon, \quad \mathcal{A} = [x, 2x] \cap \mathbb{N}, \quad \eta = \epsilon^2, \quad \nu = \epsilon^3.$$

Henceforth it will be implicit that the constants in the O and $\ll$ notation may depend on ϵ, though we will write O^* for the few occurrences where we need the constants to be absolute. Our basic idea is to count integers $mp\ell = k^2 + 1$, $k \in \mathcal{A}$. Here and elsewhere p and q always denote primes. If both m and ℓ are around Y in size, Balog's technique, if we can do the counting correctly, can show that $q \mid \ell m \Rightarrow q < Y^{\phi+\eta}$. Friedlander's idea is to make ℓ have all its prime factors $> x^{\nu}$ and so, for a fixed n, the number of solutions to $mp\ell = n^2 + 1$ is $\ll 1$. Since $pY^2 \approx x^2$, we would like p to range over values as large as possible to reduce Y, but also satisfying $p \le Y^{\phi+\eta}$.

We need to introduce a smoothing factor for the variable k in order to use a result from [3]. To this end we let $V(u)$ be an infinitely differentiable non-negative function such that

$$V(u) \begin{cases} < 2 & \text{if } 1 < u < 2, \\ = 0 & \text{if } u \le 1 \text{ or } u \ge 2, \end{cases}$$

with

$$\frac{d^r V(u)}{du^r} \ll_r 1 \quad \text{and} \quad \int_{\mathbb{R}} V(u)\, du = 1.$$

We allow implied constants to depend on the choice of $V(u)$, for example in (2) below. Since we will often have a factor $V(k/x)$, the condition $k \in \mathcal{A}$ will be superfluous and so omitted in most of the sums that follow. The following result then follows immediately from [3, Théorème 5.2] and provides us with the means of counting solutions of the required form. We write $\omega(n)$ for the number of solutions to the congruence $r^2 \equiv -1 \pmod{n}$, $0 \le r < n$, and, for $B \ge 1$, we write $b \sim B$ for $B \le b < eB$, $b \in \mathbb{N}$.

Proposition 2.1. *Let $\eta > 0$, $x, M, N \ge 1$, $MN \le x^2$, and suppose g_m, h_n are two sequences of complex numbers with modulus at most* 1. *Write*

$$r(s) = \sum_{k^2 \equiv -1 \pmod{s}} V\left(\frac{k}{x}\right) - x\frac{\omega(s)}{s}.$$

Then

$$\sum_{m \sim M} \sum_{\substack{n \sim N \\ (m,n)=1}} g_m h_n r(mn) \ll F(x, M, N, \eta), \tag{2}$$

where

$$F(x, M, N, \eta) = x^{1/2+\eta} M^{1/2} + x^{1+\eta} N^{3/2-\theta} M^{-1/4+\theta/2}. \tag{3}$$

Here $\frac{1}{4} - \theta^2$ is the best known lower bound for the smallest eigenvalue for any congruence subgroup, so $\theta = 7/64$ is acceptable by [17].

Now we will require $F(x, M, N, \eta) \ll x^{1-2\eta}$ to get an asymptotic formula for the weighted number of solutions we are counting by (2). We will take h_n as the characteristic function of the set of primes, so we will want $NM^2 \approx x^2$. Simple algebra then shows we need to take $N \approx x^{50/381}$, $M \approx x^{356/381}$. We define three sequences a_r, b_ℓ, c_s by

$$a_r = \begin{cases} 1 & \text{if } p \mid r \Rightarrow p < x^\alpha, \\ 0 & \text{otherwise,} \end{cases} \qquad c_s = 1 - a_s, \qquad b_\ell = \begin{cases} 1 & \text{if } p \mid \ell \Rightarrow p \geq x^\nu, \\ 0 & \text{otherwise.} \end{cases}$$

The Balog–Friedlander approach with this notation (omitting the smoothing factor for clarity) is to observe that

$$\sum_{\substack{k \in \mathcal{A} \\ \ell mp = k^2+1}} a_m b_\ell a_\ell \geq \Sigma_1 - \Sigma_2,$$

where

$$\Sigma_1 = \sum_{\substack{k \in \mathcal{A} \\ \ell mp = k^2+1}} b_\ell a_\ell, \qquad \Sigma_2 = \sum_{\substack{k \in \mathcal{A} \\ \ell mp = k^2+1}} b_\ell c_m.$$

Our main task will be to show that we can obtain a lower bound for Σ_1 and an upper bound for Σ_2 which are both of the "correct" size. This will lead to a lower bound of the correct order of magnitude for the integers we are counting, which includes a factor

$$\begin{aligned} 1 - 2\log\left(\frac{\log x^{\beta+\epsilon}}{\log x^\alpha}\right) &= 1 - 2\log\left(\frac{\beta+\epsilon}{\alpha}\right) = \log(1 + \epsilon/(\beta\phi)) - \log(1 + \epsilon/\beta) \\ &\approx \frac{\epsilon}{\beta}(\phi^{-1} - 1) \gg 1. \end{aligned}$$

Here we noted that $\log\phi = -1/2$ and this is what brings the $e^{-1/2}$ into the exponent of our result.

3. Preliminary results. Write $\chi(n)$ for the non-trivial character (mod 4). We note that for *all* primes $\omega(p) = 1 + \chi(p)$, and for $n \geq 2$, we have

$$\omega(p^n) = \begin{cases} 0 & \text{if } p = 2, \\ 1 + \chi(p) & \text{otherwise.} \end{cases}$$

Let $L(s, \chi)$ be the corresponding L-function. We note that, by the working in [6, Chapter 22], we have by partial summation, for $s \geq 1$,

$$\sum_{q > X} \frac{\chi(q)}{q^s} \ll \exp(-C(\log X)^{1/2}) \tag{4}$$

for some $C>0$. This result is used implicitly in the following where we need that $\omega(q)=1$ on average, and is quoted explicitly later in the proof of Theorem 1.1.

The following result is established in [8, pp. 150–151]. A more general result (with a slightly weaker error term) can be found in the works of Hooley [14] and [15, Chapter 2].

LEMMA 3.1. *For any $d\in\mathbb{N}$ and $L>1$ we have*

$$\sum_{\ell\sim L}\frac{\omega(\ell d)}{\ell}=\rho(d)\frac{L(1,\chi)}{\zeta(2)}+O\biggl(\frac{\omega(d)\tau(d)(\log x)^3}{L^{1/2}}\biggr). \tag{5}$$

Here

$$\rho(d)=\omega(d)\prod_{p\mid d}\biggl(1+\frac{1}{p}\biggr)^{-1}.$$

Now write

$$P=(x^{1-\beta-\epsilon})^2=x^{50/381-2\epsilon}.$$

The following simple lemma shows that we can add or remove the conditions $(m,p)=1$ or $(\ell,p)=1$ when counting solutions to $mp\ell=n^2+1$ with negligible error.

LEMMA 3.2. *We have, for any $\Pi\le x^{1/2}$,*

$$\sum_{x^{1/2}>p\ge\Pi}\ \sum_{\substack{n\in\mathcal{A}\\ \ell mp^2=n^2+1}}1\ll\frac{x^{1+\eta}}{\Pi}.$$

Proof. This is immediate from the well-known result that $\omega(p^2)\le 2$. ∎

Henceforth we write $B=x^{\nu}$.

LEMMA 3.3. *In the above notation, there are two sequences of reals $\lambda_d^{\pm}$ supported on the square-free numbers such that*

$$|\lambda_d^{\pm}|\le 1,\quad \lambda_d^{\pm}=0\quad \text{for } d>x^{\eta},$$

$$\sum_{d\mid n}\lambda_d^{-}\le\begin{cases}1 & \text{if } q\mid n\Rightarrow q>B,\\ 0 & \text{otherwise},\end{cases}\qquad \sum_{d\mid n}\lambda_d^{+}\ge\begin{cases}1 & \text{if } q\mid n\Rightarrow q>B,\\ 0 & \text{otherwise},\end{cases}$$

and, for λ_d equal to either of $\lambda_d^{\pm}$,

$$\sum_{d<x^{\eta}}\frac{\lambda_d\rho(d)}{d}=\bigl(1+O^*(e^{-1/\epsilon})+O((\log x)^{-1/3})\bigr)\prod_{q<B}\biggl(1-\frac{\rho(q)}{q}\biggr). \tag{6}$$

Proof. See [16, Lemma 3]. This is a "Fundamental Lemma" form of the result (since we are sieving by the primes up to x^{ν} with distribution level x^{η} where $\eta/\nu=\epsilon^{-1}$ is "large"). ∎

We will need a less precise form of the above result for a narrow range of the variables which again follows from [16, Lemma 3] as a simple upper bound. Henceforth γ always denotes Euler's constant.

LEMMA 3.4. *There is a sequence of reals λ'_d supported on the square-free numbers such that*

$$|\lambda'_d| \le 1, \quad \lambda'_d = 0 \quad \text{for } d > B, \qquad \sum_{d \mid n} \lambda'_d \ge \begin{cases} 1 & \text{if } q \mid n \Rightarrow q > B, \\ 0 & \text{otherwise}, \end{cases}$$

and

$$\sum_{d<B} \frac{\lambda'_d \rho(d)}{d} \le 2e^{\gamma}\big(1 + O((\log x)^{-1/3})\big) \prod_{q<B} \left(1 - \frac{\rho(q)}{q}\right). \tag{7}$$

LEMMA 3.5. *We have*

$$\prod_{q<B} \left(1 - \frac{\rho(q)}{q}\right) = \frac{e^{-\gamma}}{\nu \log x} \frac{\zeta(2)}{L(1,\chi)} \left(1 + O\left(\frac{1}{\log x}\right)\right). \tag{8}$$

Proof. This is essentially [13, Lemma 3.2] and also occurs inter alia in [5, p. 10], and [7, p. 10]. ∎

In the final part of our proof of Theorem 1.1 we shall not be able to make use of the averaging over ℓ given in Lemma 3.1. This forces us to consider

$$\sum_{d \le x^{\eta}} \frac{\lambda_d}{d} \omega(md) = \frac{\omega(m)}{m} \sum_{\substack{d \le x^{\eta} \\ 2 \nmid (d,m)}} \frac{\lambda_d}{d} \omega\left(\frac{d}{(d,m)}\right). \tag{9}$$

Thus, in our sieve bounds,

$$\prod_{q<B} \left(1 - \frac{\rho(q)}{q}\right)$$

is replaced by (for $\omega(m) \ne 0$)

$$\prod_{\substack{q<B \\ q \nmid m}} \left(1 - \frac{\omega(q)}{q}\right) \prod_{\substack{2<q<B \\ q \mid m}} \left(1 - \frac{1}{q}\right).$$

For this reason we introduce

$$\Omega(m) = \prod_{\substack{q \mid m \\ q>2}} \frac{1 - 1/q}{1 - 2/q} = \prod_{\substack{q \mid m \\ q>2}} \left(1 + \frac{1}{q-2}\right).$$

We then need the following result.

LEMMA 3.6. *For all $Y \ge 2$ we have*

$$\sum_{m \le Y} \frac{\omega(m)\Omega(m)}{m} \ll \log Y. \tag{10}$$

Proof. We have

$$\sum_{m\le Y}\frac{\omega(m)\Omega(m)}{m}\le\prod_{q\le Y}\left(1+\frac{\omega(q)\Omega(q)}{q}+\frac{\omega(q^2)\Omega(q^2)}{q^2}+\cdots\right)$$
$$=\frac{3}{2}\prod_{\substack{2<q\le Y\\ \omega(q)\ne 0}}\left(1+\frac{2}{q}+O\left(\frac{1}{q^2}\right)\right)\ll\log Y,$$

by standard procedures (we are, of course, using (4) implicitly here). ∎

4. Proof of Theorem 1.1. Pick $\Pi=e^hP$ with $1\le e^h<x^\eta$. Then choose $L=e^gx^{\beta+\epsilon-\eta}$ with $1\le e^g<x^\eta$. We first want to find a lower bound very close to the expected formula for

$$S(L,\Pi)=\sum_{\substack{k\in\mathcal{A}\\ \ell mp=k^2+1\\ \ell\sim L,\,p\sim\Pi}}b_\ell a_\ell V\left(\frac{k}{x}\right).$$

By Proposition 2.1 and Lemma 3.2 we have

$$S(L,\Pi)=x\sum_{\ell\sim L,\,p\sim\Pi}b_\ell a_\ell\frac{\omega(p)\omega(\ell)}{p\ell}+O(x^{1-2\eta}).$$

Now

$$a_\ell=1-\sum_{\substack{q\mid\ell\\ x^\alpha<q<eL}}1.$$

Let

$$Q_1=x^\alpha,\qquad Q_2=L(\log x)^{-20},\qquad Q_3=eL.$$

For $1\le j\le 2$ write $\mathcal{Q}_j=[Q_j,Q_{j+1})\cap\mathbb{N}$. We then write

$$a_\ell=1-\sum_{j=1}^{2}\sigma_j(\ell)\quad\text{where}\quad\sigma_j(\ell)=\sum_{\substack{q\mid\ell\\ q\in\mathcal{Q}_j}}1.$$

We then have three terms to deal with in order to evaluate $S(L,\Pi)$ as follows:

(i) The first term is

$$\sum_{\ell\sim L,\,p\sim\Pi}b_\ell\frac{\omega(p)\omega(\ell)}{p\ell}=\sum_{\ell\sim L}b_\ell\frac{\omega(\ell)}{\ell}\sum_{p\sim\Pi}\frac{\omega(p)}{p}.$$

Now

$$\sum_{p\sim\Pi}\frac{\omega(p)}{p}=\frac{1}{\log\Pi}\big(1+O((\log x)^{-1})\big),$$

while from Lemma 3.1,

$$\sum_{\ell\sim L} b_\ell \frac{\omega(\ell)}{\ell} \geq \sum_{\ell\sim L} \frac{\omega(\ell)}{\ell} \sum_{d|\ell} \lambda_d^- \\ = \frac{L(1,\chi)}{\zeta(2)} \sum_{d<x^\eta} \frac{\lambda_d^- \rho(d)}{d} + O\Big(\sum_{d<x^\eta} \frac{\omega(d)\tau(d)(\log x)^3}{dL^{1/2}}\Big).$$

The error term above is clearly $\ll (\log x)^7 L^{-1/2}$. By Lemma 3.3 the main term in the last line is

$$\frac{L(1,\chi)}{\zeta(2)} K(x,\epsilon) \prod_{q<B}\Big(1-\frac{\rho(q)}{q}\Big) = K(x,\epsilon)\frac{e^{-\gamma}}{\nu\log x}\Big(1+O\Big(\frac{1}{\log x}\Big)\Big),$$

using Lemma 3.5. Here we have written

$$K(x,\epsilon) = \big(1+O^*(\eta)+O((\log x)^{-1/3})\big),$$

noting that $\exp(-1/\epsilon) \ll \epsilon^2 = \eta$. Combining all our results gives

$$\sum_{\ell\sim L,\, p\sim \Pi} b_\ell \frac{\omega(p)\omega(\ell)}{p\ell} \geq \frac{e^{-\gamma}}{(\log \Pi)(\nu \log x)} K(x,\epsilon).$$

(ii) We have

$$\sum_{\ell\sim L} b_\ell \sigma_1(\ell)\frac{\omega(\ell)}{\ell} = \sum_{q\in\mathcal{Q}_1} \frac{\omega(q)}{q} \sum_{\ell\sim L/q} b_\ell \frac{\omega(\ell)}{\ell}.$$

We can treat

$$\sum_{\ell\sim L/q} b_\ell \frac{\omega(\ell)}{\ell}$$

as in case (i) except that we now require an upper bound. We thus switch λ^- to λ^+ and the error term

$$O\big((\log x)^7 L^{-1/2}\big) \quad \text{becomes} \quad O\big((\log x)^7 (L/q)^{-1/2}\big).$$

Since $q < Q_2 = L(\log x)^{-20}$, this error term is still admissible. (In fact, we must have $q < eL/B$ as ℓ has no prime factors $< B$, but we are only obtaining an upper bound, and this situation does not arise in the analogous case when we switch the rôles of ℓ and m.) We deduce from (4) that

$$\begin{aligned}\sum_{q\in\mathcal{Q}_1}\frac{\omega(q)}{q} &= \sum_{q\in\mathcal{Q}_1}\frac{1}{q}+O\big(\exp(-C(\log x)^{1/2})\big)\\ &= \log\left(\frac{\log L-20\log\log x}{\log x^{\alpha}}\right)+O((\log x)^{-1})\\ &= \log\left(\frac{\frac{356}{381}+\epsilon}{\phi\frac{356}{381}+\epsilon}\right)+O^*(\eta)+O((\log x)^{-1/2})\\ &= \frac{1}{2}-R\epsilon+O^*(\eta)+O((\log x)^{-1/2})\end{aligned}$$

where

$$R=\frac{381}{356}(e^{1/2}-1)>\frac{1}{2}.$$

Hence

$$\sum_{\ell\sim L,\,p\sim\Pi} b_\ell\frac{\omega(p)\omega(\ell)}{p\ell}\sigma_1(\ell)\le\frac{e^{-\gamma}}{(\log\Pi)(\nu\log x)}K(x,\epsilon)\left(\frac{1}{2}-R\epsilon\right).$$

(iii) In this case for large x we must have $q>eL/B$. This forces $q\sim L$ as ℓ cannot have a prime factor less than B, so $\ell=1$. The contribution from this final part of the sum is therefore

$$\sum_{q\sim L}\frac{\omega(q)}{q}\frac{1}{\log\Pi}=\frac{1+O((\log x)^{-1})}{\log\Pi\log L}=\frac{K(x,\epsilon)O^*(\nu)}{(\log\Pi)(\nu\log x)}.$$

We have thus established that

$$S(L,\Pi)\ge x\frac{1+\epsilon}{2}\frac{e^{-\gamma}}{(\log\Pi)(\nu\log x)}K(x,\epsilon).$$

It follows that (recall $0\le g,h\le\eta\log x$)

$$\sum_{g,h}S(L,\Pi)\ge\eta^2x\frac{1+\epsilon}{2}\frac{e^{-\gamma}}{\left(\frac{50}{381}-2\epsilon\right)\nu}K(x,\epsilon). \tag{11}$$

We must now change the rôles of the variables to estimate the quantity we called Σ_2 in §2. Instead of breaking up the summation range over ℓ we must do this for m. We treat p as before and suppose that $\Pi=e^hP$ with $e^h<x^\eta$. To ensure we include all possible values for m (since we are subtracting the final term, we need an upper bound) we consider

$$\frac{x^{2-\beta-\epsilon}}{e^2\Pi}<m<\frac{e^2x^{2-\beta-\epsilon+\eta}}{\Pi}.$$

So we will be taking values

$$M=\frac{e^gx^{2-\beta-\epsilon}}{e^2\Pi}\quad\text{with}\quad e^g<e^4x^\eta.$$

We are thus summing over $\eta \log x + O(1)$ values M and the additional $O(1)$ introduces no difficulties here. We must therefore study

$$T(M,\mathit{\Pi}) = \sum_{\substack{\ell mp = k^2+1 \\ m\sim M,\, p\sim \mathit{\Pi}}} b_\ell c_m V\left(\frac{k}{x}\right).$$

By Lemma 3.3 we can give an upper bound for this quantity by considering

$$\sum_{d<x^\eta} \lambda_d^+ \sum_{\substack{\ell dmp = k^2+1 \\ m\sim M,\, p\sim \mathit{\Pi}}} c_m V\left(\frac{k}{x}\right).$$

We apply Proposition 2.1 to demonstrate this sum is

$$x \sum_{d<x^\eta} \lambda_d^+ \sum_{\substack{p\sim \mathit{\Pi} \\ m\sim M}} c_m \frac{\omega(p)\omega(md)}{pdm} + O(x^{1-\eta}).$$

Now

$$c_m = \sum_{\substack{q\mid m \\ x^\alpha < q < eM}} 1 = c_m(1) + c_m(2), \quad \text{say},$$

where $q < eM(\log x)^{-20}$ in $c_m(1)$. We can work as in case (ii) of the estimate for $S(L,\mathit{\Pi})$ to obtain a satisfactory bound for this part of the sum, namely,

$$\le \frac{e^{-\gamma} x}{(\log \mathit{\Pi})(\nu \log x)} K(x,\epsilon)\left(\frac{1}{2} - R\epsilon\right).$$

The sum involving $c_m(2)$ is

$$x \sum_{m\le (\log x)^{20}} \sum_{d<x^\eta} \lambda_d^+ \sum_{\substack{p\sim \mathit{\Pi} \\ mq\sim M}} \frac{\omega(p)\omega(mdq)}{pdmq}. \tag{12}$$

Of course $\omega(mdq) = \omega(q)\omega(dm)$. Clearly

$$\sum_{qm\sim M} \frac{\omega(q)}{q} = \frac{1}{\log(M/m)} + O((\log M)^{-2}).$$

We then use (9) and the working that follows in §3 to get the contribution from (12) to be

$$\ll \frac{x \log\log x}{(\log x)^3}.$$

We have thus shown that

$$T(M,N) \le \frac{e^{-\gamma} x}{(\log \mathit{\Pi})(\nu \log x)} K(x,\epsilon)\left(\frac{1}{2} - R\epsilon\right) + O\left(\frac{x \log\log x}{(\log x)^3}\right).$$

It follows that (recall $0 \le g \le 4 + \eta \log x$, and $0 \le h \le \eta \log x$)

$$\sum_{g,h} T(M, \Pi) \le \eta^2 x \frac{1-\epsilon}{2} \frac{e^{-\gamma}}{\left(\frac{50}{381} - 2\epsilon\right)\nu} K(x, \epsilon). \tag{13}$$

Taking the difference between (11) and (13) gives a lower bound for the numbers we wish to count, which is

$$\ge \epsilon\eta^2 x \frac{e^{-\gamma}}{\left(\frac{50}{381} - 2\epsilon\right)\nu}\left(1 + O^*(\epsilon) + O((\log x)^{-1/3})\right).$$

This completes the proof of Theorem 1.1.

5. Primitive divisors of quadratic polynomials. We recall the following standard definition and proposition (see, for example, [8, 9]).

Definition 5.1. Let (A_n) denote a sequence with integer terms. We say an integer $d > 1$ is a *primitive divisor* of A_n if

(1) $d \mid A_n$,
(2) $\gcd(d, A_m) = 1$ for all non-zero terms A_m with $m < n$.

Proposition 5.2. *For all $n > |D|$, the term $P_n = n^2 + D$ has a primitive divisor if and only if $P^+(n^2+D) > 2n$. For all $n > |D|$, if P_n has a primitive divisor then that primitive divisor is a prime and it is unique.*

In [8] we proved the following result (we take $D = 1$ for simplicity, but as with our previous sections the results can be made more general).

Theorem 5.3. *Define*

$$\rho(x) = |\{n \le x : n^2 + 1 \text{ has a primitive divisor}\}|.$$

For all sufficiently large x we have

$$0.5324 < \frac{\rho(x)}{x} < 0.905.$$

We also tentatively suggested the following conjecture.

Conjecture 5.4. As $x \to \infty$ we have $\rho(x) \sim x \log 2$.

It was explained there that such a conjecture would imply astonishingly strong results on the lower bound for $P^+(n^2+1)$ for infinitely many n. Since this looks unlikely to be established without a significant advance in knowledge, it seems worthwhile to give a modest sharpening of Theorem 5.3:

Theorem 5.5. *For all sufficiently large x we have*

$$0.5377 < \frac{\rho(x)}{x} < 0.86. \tag{14}$$

Proof. To consider the upper bound in our result we need to use the working in [19], or rather the working with one factor changed.

Let $P_x = x^\tau = \max_{n \le x} P^+(n^2+1)$. The basic argument goes back to Chebyshev and starts with the observation that

$$\sum_{x \le p \le P_x} G_p \log p = x \log x + O(x).$$

Here

$$G_p = \sum_{p \mid k^2+1} V\left(\frac{k}{x}\right).$$

Simplifying a few details to expose the main idea, we then wish to obtain upper bounds for G_p of the form

$$\sum_{1 \le p x^{-\alpha} \le e} G_p \le K(\alpha)(1+o(1)) \frac{X}{\log x}.$$

It is then a question of showing that

$$\int_1^\tau \alpha K(\alpha)\, d\alpha < 1,$$

and this determines the maximum value for τ. In [8] we show that in the above notation

$$\frac{\rho(x)}{x} \le (1+o(1)) \int_1^\tau K(\alpha)\, d\alpha.$$

So, to prove our result, we only need to perform the same calculations as in [19], removing the factor α from the integrand. In some ranges of α the integrals are elementary, but in others numerical integration must be employed. In his paper [19, p. 1268] Merikoski has kindly supplied links to his Python programs for these calculations. Changing these programs to remove the α factor, and calculating the remaining elementary integrals, then gives

$$\int_1^\tau K(\alpha)\, d\alpha < 0.86$$

as required to complete the proof.

We give one example of the integrals and calculations involved to illustrate what happens. In [19] the argument splits according to the size of α and the first range is $1 \le \alpha < 758/733$. The integral computed for this region is

$$\int_1^{758/733} 1\, d\alpha + G_1 = 0.034106\ldots + G_1,$$

where

$$G_1 := \int_1^{758/733} \alpha \left(\int_{\sigma}^{\alpha - 2\sigma} \omega(\alpha/\beta - 1) \frac{d\beta}{\beta^2} + \int_{\xi}^{\alpha/2} \omega(\alpha/\beta - 1) \frac{d\beta}{\beta^2} \right) d\alpha$$
$$< 0.01745.$$

Here $\omega(u)$ is Buchstab's function, and σ, ξ are certain functions of α given in [19, §2.4.1].

For the proof of our result the contribution from that region is

$$\int_1^{758/733} \frac{1}{\alpha}\, d\alpha + H_1 = 0.033537\ldots + H_1$$

where

$$H_1 := \int_1^{758/733} \left(\int_{\sigma}^{\alpha - 2\sigma} \omega(\alpha/\beta - 1) \frac{d\beta}{\beta^2} + \int_{\xi}^{\alpha/2} \omega(\alpha/\beta - 1) \frac{d\beta}{\beta^2} \right) d\alpha$$
$$< 0.01706.$$

Of course, α is still quite close to 1 in this region, so dividing by α has only made a small change here.

To consider the lower bound in (14) we note that the working in [8] shows that

$$\frac{\rho(x)}{x} \geq (1 + o(1)) \int_{\tau}^{2} K(\alpha)\, d\alpha, \tag{15}$$

where τ is the solution to

$$\int_{\tau}^{2} \alpha K(\alpha)\, d\alpha = 1.$$

In [8] we used an elementary argument to allow the choice

$$K(\alpha) = \frac{2}{\alpha - 1}.$$

If, instead, we use Proposition 2.1 (now with a different smoothing factor $V(k/x)$ providing an upper bound which only loses an η factor in the main term) with $M = x^{1-4\eta}$ (compare [3, Théorème 1.1]), we can replace this with (ignoring an η term for clarity)

$$K(\alpha) = \frac{2}{\alpha - \frac{153}{178}}.$$

Calculations then give $\tau = 1.73111\ldots$, leading, via (15), to the lower bound in (14). We note that the corresponding value of τ in [8] was $1.766249\ldots$.

This change may appear quite small, but subtracting both values from 2 (which is the "starting point", so to speak), this is a 15% improvement. ■

Acknowledgements. The author thanks the referee for their careful reading of the paper and pertinent remarks.

References

[1] A. Balog, *$p + a$ without large prime factors*, in: Sém. Théor. Nombres Bordeaux 1983–1984, Univ. Bordeaux I, 1984, exp. no. 31, 5 pp.

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

[3] R. de la Bretèche et S. Drappeau, *Niveau de repártition des polynômes quadratiques et crible majorant pour les entiers friables*, J. Eur. Math. Soc. 22 (2020), 1577–1624.

[4] J. W. Bober, D. Fretwell, G. Martin and T. D. Wooley, *Smooth values of polynomials*, J. Austral. Math. Soc. 108 (2020), 245–261.

[5] C. Dartyge, *Entiers de la forme $n^2 + 1$ sans grand facteur premier*, Acta Math. Hungar. 72 (1996), 1–34.

[6] H. Davenport, *Multiplicative Number Theory*, 2nd ed., Grad. Texts in Math. 74, Springer, 1980.

[7] J.-M. Deshouillers and H. Iwaniec, *On the greatest prime factor of $n^2 + 1$*, Ann. Inst. Fourier (Grenoble) 32 (1982), no. 4, 1–11.

[8] G. R. Everest and G. Harman, *On primitive divisors of $n^2 + b$*, in: Number Theory and Polynomials, London Math. Soc. Lecture Note Ser. 352, Cambridge Univ. Press, 2008, 142–154.

[9] G. R. Everest, S. Stevens, D. Tamsett and T. Ward, *Primes generated by recurrence sequences*, Amer. Math. Monthly 114 (2007), 417–431.

[10] J. B. Friedlander, *Shifted primes without large prime factors*, in: Number Theory and Applications, Kluwer, 1989, 393–401.

[11] G. H. Hardy and J. E. Littlewood, *Some problems of 'partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. 44 (1923), 1–70.

[12] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th ed., Oxford Univ. Press, 2008.

[13] G. Harman, *On values of $n^2 + 1$ free of large prime factors*, Arch. Math. (Basel) 90 (2008), 239–245.

[14] C. Hooley, *On the greatest prime factor of a quadratic polynomial*, Acta Math. 117 (1967), 281–299.

[15] C. Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge Tracts in Math. 70, Cambridge Univ. Press, 1976.

[16] H. Iwaniec, *A new form of the error term in the linear sieve*, Acta Arith. 37 (1980), 307–320.

[17] H. H. Kim, *Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2*, J. Amer. Math. Soc. 16 (2003), 139–183.

[18] G. Martin, *An asymptotic formula for the number of smooth values of a polynomial*, J. Number Theory 93 (2002), 108–182.

[19] J. Merikoski, *On the largest prime factor of $n^2 + 1$*, J. Eur. Math. Soc. 25 (2023), 1253–1284.

[20] A. Schinzel, *On two theorems of Gelfond and some of their applications*, Acta Arith. 13 (1967), 177–236.

Glyn Harman
Department of Mathematics
Royal Holloway, University of London
Egham, Surrey TW20 0EX, UK
E-mail: g.harman@rhul.ac.uk