Mathematika 61 (2015) 741–755 © 2014 University College London
doi:10.1112/S002557931400031X

ALMOST-PRIME VALUES OF BINARY FORMS WITH ONE PRIME VARIABLE

A. J. IRVING

Abstract. By establishing an improved level of distribution we study almost-primes of the form $f(p, n)$ where f is an irreducible binary form over $\mathbb{Z}$.

§1. *Introduction.* A well known problem in number theory is to show that if $f \in \mathbb{Z}[x]$ is an irreducible polynomial with $\deg f \geqslant 2$ then, provided the values of f have no fixed prime divisor, there are infinitely many $n \in \mathbb{Z}$ for which $f(n)$ is prime. This seems to be out of reach of current methods. However, using sieves one can show that there are infinitely many $n \in \mathbb{Z}$ for which $f(n)$ has a small number of prime factors. Let P_r denote numbers with at most r prime factors, counted with multiplicities, and let $k = \deg f$. Richert [7] showed that there are infinitely many n for which $f(n)$ is a P_{k+1}. An even harder question is to ask whether there are infinitely many primes p for which $f(p)$ is itself prime. This was also considered by Richert, who showed that there are infinitely many p for which $f(p)$ is a P_{2k+1} (provided we impose conditions on f to avoid the obvious counterexamples).

Both problems are made easier if we consider irreducible binary forms $f \in \mathbb{Z}[x, y]$ instead of single variable polynomials. A theorem of Fermat states that any prime $p \equiv 1 \pmod 4$ is the sum of two squares and therefore the binary quadratic form $m^2 + n^2$ represents infinitely many primes. The case of a general binary quadratic form was handled by Dirichlet. Much more recently, Heath-Brown [5] showed that the cubic $m^3 + 2n^3$ represents infinitely many primes. If f is a binary form with $k \geqslant 4$ then the best result known is due to Greaves [3], who showed that if f is irreducible then the values $f(m, n)$ are infinitely often $P_{[k/2]+1}$, provided of course that they have no fixed prime divisor. In this paper we will consider the values $f(p, n)$ of a binary form where n is an integer and p a prime. A result of Fouvry and Iwaniec [2] shows that there are infinitely many primes of the form $p^2 + n^2$; we are unaware of any existing results dealing with higher degree forms. It is clear that by fixing the prime variable p and applying the above result of Richert to the resulting polynomial values we can obtain infinitely many P_{k+1}. We will improve this result for all $k \geqslant 3$ as follows.

THEOREM 1.1. *Let $f \in \mathbb{Z}[x, y]$ be an irreducible binary form of degree $k \geqslant 3$. Suppose that for every prime p we have*

$$\#\{n \pmod p : f(1, n) \equiv 0 \pmod p\} < p.$$

Received 9 May 2014, published online 5 December 2014.
MSC (2010): 11N32, 11N36 (primary).

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

There are then infinitely many pairs (p, n) with $n \in \mathbb{Z}$ and p prime for which $f(p, n)$ is a $P_{[3k/4]+1}$.

The proof of this depends on an improved "level of distribution" result for the values $f(p, n)$. Roughly speaking, we count the number of these which are divisible by an integer d when p and n have size N. If we were to consider each prime p separately then we could only handle $d \leqslant N^{1-\delta}$ for any $\delta > 0$. We will show that we can obtain a result on average over d provided that $d \leqslant N^{4/3-\delta}$. Theorem 1.1 then follows easily by using the weighted sieve. The details of our level of distribution are somewhat technical so we will leave a precise statement until §3.

Our level of distribution should be compared with Fouvry and Iwaniec's for the values $p^2 + n^2$ [2, Lemma 4]. In our notation their result essentially states that one can take d as large as $N^{2-\delta}$ for that form. Their proof depends crucially on the fact that the roots of the congruence $n^2 + 1 \equiv 0 \pmod{d}$ satisfy very strong distribution properties. This enables them to prove a large sieve inequality for the fractions n/d which is essentially optimal. Our result also depends on a large sieve type inequality. However we do not have comparable distribution estimates for the roots of higher degree polynomial congruences and therefore our level of distribution is weaker. In the next section we will give details of the variant of the large sieve we use. It concerns the sum of a sequence of coefficients α_m, for example the indicator function of the primes, over the points (m, n) in a sublattice of $\mathbb{Z}^2$. We will show that if we average over a suitable family of lattices then we can control such a sum. To reduce the binary form question to one concerning lattices we use methods similar to those of Daniel [1].

Throughout this paper we use the notation $(a; b)$ for the highest common factor of the integers a and b. We write $x \sim y$ for the inequality $y \leqslant x < 2y$. The notation $\|x\|$ denotes the Euclidean length of a vector $x \in \mathbb{R}^2$. We will denote the indicator function of the primes by $\mathbf{1}_P(n)$. We fix a smooth function W which has compact support in $[0, 1]$ and which takes nonnegative values. Finally we adopt the standard convention that ϵ denotes a small positive quantity whose value may differ at each occurrence. All our implied constants may depend on ϵ, W and the binary form f.

§2. *A large sieve for lattices.*

2.1. *Introduction.* Let α_m be a sequence of complex numbers with $|\alpha_m| \leqslant 1$ and let $\lambda \subseteq \mathbb{Z}^2$ be a lattice. For $N \geqslant 0$ we are interested in the quantity

$$\psi(\lambda, N, \alpha) = \sum_{(m,n) \in \lambda \cap (0,N] \times \mathbb{Z}} \alpha_m W\left(\frac{n}{N}\right).$$

We expect that for a typical λ we have

$$\psi(\lambda, N, \alpha) \approx \frac{N \hat{W}(0)}{\det \lambda} \sum_{m \leqslant N} \alpha_m.$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

We will show that this holds if we average over a suitable set of lattices λ. We will only consider the case that the set of m-coordinates of points in λ

$$\{m : (m, n) \in \lambda\}$$

has greatest common factor 1, since if this does not hold then only a homogeneous arithmetic progression of m occurs so the result cannot be true.

We will write $\det \lambda = d$ and restrict our consideration to lattices with $d \sim D$ for some parameter D. For a given lattice λ we let B_1 be a nonzero element of λ of minimal length and B_2 be a vector of minimal length in the elements of λ which are not multiples of B_1. It is well known that $\{B_1, B_2\}$ is a basis for λ and that

$$\|B_1\|\|B_2\| \asymp \det \lambda.$$

Let B be the matrix with rows B_1, B_2. Since we are free to choose the signs of both B_1 and B_2 we may assume that $B_{11} \geqslant 0$ and $\det B = \det \lambda$. We know that $\|B_1\| \ll (\det \lambda)^{1/2}$ and thus we have the same bound for B_{11} and B_{12}. We will consider an average over lattices where each possible value for B_{11} occurs at most once but we make no assumption on the distribution of the remaining entries in B. Our result is then as follows. It should be noted that the shortest nonzero vector in λ may not be unique. In this case we are free to choose the vector in such a way that the conditions of the theorem are satisfied.

THEOREM 2.1. *Let α_m be a sequence of complex numbers with $|\alpha_m| \leqslant 1$ and let $D, M_1 \geqslant 1$. Let Λ be a set of lattices in $\mathbb{Z}^2$ such that if $\lambda \in \Lambda$ then $\det \lambda \sim D$ and, letting B be as above, we have $B_{11} \sim M_1$. Assume that for each $\lambda \in \Lambda$ the m-coordinates of points are coprime, as described above. In addition, suppose that for each $m \sim M_1$ we have*

$$\#\{\lambda \in \Lambda : B_{11}(\lambda) = m\} \leqslant 1.$$

Suppose that $\delta > 0$.

(1) *If $D \leqslant N^{1-\delta}$ then for any $A > 0$ we have*

$$\sum_{\lambda \in \Lambda} \left| \psi(\lambda, N, \alpha) - \frac{N \hat{W}(0)}{\det \lambda} \sum_{m \leqslant N} \alpha_m \right| \ll_{\delta, A} N^{-A}.$$

(2) *If*

$$N^{1-\delta} \leqslant D < M_1 N^{1-\delta}$$

then

$$\sum_{\lambda \in \Lambda} \left| \psi(\lambda, N, \alpha) - \frac{N \hat{W}(0)}{\det \lambda} \sum_{m \leqslant N} \alpha_m \right| \ll_{\epsilon, \delta} N^{1+2\delta+\epsilon} M_1^{-1/2} D^{1/2}$$

for any $\epsilon > 0$.

It is useful to know when this result is nontrivial. We note that, since $\#\Lambda \ll M_1$, we have

$$\sum_{\lambda \in \Lambda} \frac{N \hat{W}(0)}{\det \lambda} \sum_{m \leqslant N} \alpha_m \ll \frac{N^2 M_1}{D}$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

and that

$$N^{1+2\delta+\epsilon}M_1^{-1/2}D^{1/2} < \frac{N^2M_1}{D}$$

if and only if

$$D < N^{2/3-4\delta/3-2\epsilon/3}M_1.$$

Our bound can therefore only be nontrivial if $D \leqslant N^{2/3-\eta}M_1$ for some $\eta > 0$. In particular, since $M_1 \ll D^{1/2}$ the largest D we can handle is $D \ll N^{4/3-\eta}$. However, if M_1 is smaller then the range of D must be decreased.

2.2. *Transforming the sum.* We can write

$$\lambda = \{(u, v)B : (u, v) \in \mathbb{Z}^2\}.$$

Our assumption that the m-coordinates of points in λ have greatest common factor 1 implies that we must have $(B_{11}; B_{21}) = 1$. In addition, since $B_{11} \sim M_1 \geqslant 1$ we have $B_{11} > 0$.

For a fixed $m \in (0, N]$ we consider the quantity

$$S(m) = \sum_{\substack{n\in\mathbb{Z}\\(m,n)\in\lambda}} W\left(\frac{n}{N}\right) = \sum_{\substack{(u,v)\in\mathbb{Z}^2\\B_{11}u+B_{21}v=m}} W\left(\frac{B_{12}u + B_{22}v}{N}\right).$$

The condition

$$m = B_{11}u + B_{21}v$$

is equivalent to

$$m \equiv B_{21}v \pmod{B_{11}}$$

in which case

$$u = \frac{m - B_{21}v}{B_{11}}.$$

We therefore have

$$\begin{aligned} S(m) &= \sum_{v\equiv m\overline{B_{21}} \pmod{B_{11}}} W\left(\frac{B_{12}(m - B_{21}v) + B_{11}B_{22}v}{B_{11}N}\right) \\ &= \sum_{v\equiv m\overline{B_{21}} \pmod{B_{11}}} W\left(\frac{B_{12}m + dv}{B_{11}N}\right) \\ &= \sum_{u\in\mathbb{Z}} W\left(\frac{B_{12}m + d(m\overline{B_{21}} + uB_{11})}{B_{11}N}\right) \\ &= \sum_{u\in\mathbb{Z}} W\left(\frac{m(B_{12} + d\overline{B_{21}})}{B_{11}N} + \frac{du}{N}\right). \end{aligned}$$

We may now apply the Poisson summation formula to deduce that

$$S(m) = \frac{N}{d}\sum_{v\in\mathbb{Z}} \hat{W}\left(\frac{vN}{d}\right)e\left(\frac{mv(B_{12} + d\overline{B_{21}})}{dB_{11}}\right).$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

We therefore conclude that

$$\psi(\lambda, N, \alpha) = \frac{N}{d} \sum_{v \in \mathbb{Z}} \hat{W}\left(\frac{vN}{d}\right) \sum_{m \leqslant N} \alpha_m e\left(\frac{mv(B_{12} + d\overline{B_{21}})}{dB_{11}}\right).$$

The $v = 0$ term in this is

$$\frac{N\hat{W}(0)}{d} \sum_{m \leqslant N} \alpha_m,$$

which is precisely the main term we require.

For any $A \in \mathbb{N}$ we may integrate by parts A times to obtain the standard estimate

$$\hat{W}(x) \ll_A \min(1, |x|^{-A}).$$

Recall that we have $d \sim D$. We will truncate the sum over v to $|v| \leqslant DN^{-1+\delta}$. Specifically, for any $\delta > 0$ and $A \in \mathbb{N}$ we have

$$\frac{N}{d} \sum_{|v| > DN^{-1+\delta}} \hat{W}\left(\frac{vN}{d}\right) \sum_{m \leqslant N} \alpha_m e\left(\frac{mv(B_{12} + d\overline{B_{21}})}{dB_{11}}\right) \ll_{\delta, A} N^{-A}.$$

Combining all of the above we see that

$$\psi(\lambda, N, \alpha) = \frac{N\hat{W}(0)}{d} \sum_{m \leqslant N} \alpha_m + \psi_1(\lambda, N, \alpha, \delta) + O_{\delta, A}(N^{-A})$$

where

$$\psi_1(\lambda, N, \alpha, \delta) = \frac{N}{d} \sum_{0 < |v| \leqslant DN^{-1+\delta}} \hat{W}\left(\frac{vN}{d}\right) \sum_{m \leqslant N} \alpha_m e\left(\frac{mv(B_{12} + d\overline{B_{21}})}{dB_{11}}\right).$$

It remains to bound ψ_1, at least on average over λ. This is trivial if $DN^{-1+\delta} < 1$, that is $D < N^{1-\delta}$, as then $\psi_1 = 0$. This is thus enough to prove the first assertion in Theorem 2.1. We may therefore assume that $D \geqslant N^{1-\delta}$.

We have

$$\psi_1(\lambda, N, \alpha, \delta) \ll \frac{N}{D} \sum_{0 < |v| \leqslant DN^{-1+\delta}} \left| \sum_{m \leqslant N} \alpha_m e\left(\frac{mv(B_{12} + d\overline{B_{21}})}{dB_{11}}\right)\right|.$$

We will remove the factor $e(mvB_{12}/dB_{11})$ using partial summation. This results in

$$\psi_1(\lambda, N, \alpha, \delta) \ll \frac{N}{D}\left(1 + N^{\delta}\frac{|B_{12}|}{B_{11}}\right) \sum_{0 < |v| \leqslant DN^{-1+\delta}} \max_{N' \leqslant N} \left| \sum_{m \leqslant N'} \alpha_m e\left(\frac{mv\overline{B_{21}}}{B_{11}}\right)\right|.$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

Recalling that $B_{12} \ll D^{1/2}$, $B_{11} \sim M_1 \ll D^{1/2}$ and using our assumption that each B_{11} occurs at most once we thus see that

$$\sum_{\lambda\in\Lambda} |\psi_1(\lambda, N, \alpha, \delta)| \ll N^{1+\delta} D^{-1/2} M_1^{-1} \times \sum_{B_{11}\sim M_1} \max_{(B_{21}; B_{11})=1} \sum_{0<|v|\leqslant DN^{-1+\delta}} \max_{N'\leqslant N} \left| \sum_{m\leqslant N'} \alpha_m e\left(\frac{mv\overline{B_{21}}}{B_{11}}\right)\right|.$$

By Cauchy's inequality we may bound this by

$$N^{1/2+3\delta/2} M_1^{-1/2} \psi_2(\Lambda, N, \alpha, \delta)^{1/2}$$

where

$$\psi_2(\Lambda, N, \alpha, \delta) = \sum_{B_{11}\sim M_1} \max_{(b; B_{11})=1} \sum_{0<|v|\leqslant DN^{-1+\delta}} \max_{N'\leqslant N} \left| \sum_{m\leqslant N'} \alpha_m e\left(\frac{mvb}{B_{11}}\right)\right|^2.$$

2.3. *Applying the large sieve.* Each vb/B_{11} occurring in ψ_2 is congruent mod $\mathbb{Z}$ to a unique a/q with $(a; q) = 1$, $0 \leqslant a < q$ and $q \ll M_1$. We will group together terms with the same a/q and bound the resulting sums over dyadic intervals $q \sim Q$. We must therefore give an upper bound for the number of times each a/q occurs in our sum.

LEMMA 2.2. *Assume that $D, M_1 \geqslant 1$ and $\delta > 0$ satisfy*

$$N^{1-\delta} \leqslant D < M_1 N^{1-\delta}.$$

Suppose that for each integer $B_{11} \sim M_1$ we are given an integer b with $(b; B_{11}) = 1$. Then, if $(a; q) = 1$ and $0 \leqslant a < q \ll M_1$, we have

$$\#\left\{B_{11} \sim M_1, 0 < |v| \leqslant DN^{-1+\delta} : \frac{vb}{B_{11}} \equiv \frac{a}{q} \pmod{\mathbb{Z}}\right\} = \begin{cases} 0 & q < N^{1-\delta} M_1 D^{-1} \\ O(M_1 q^{-1}) & \text{otherwise.} \end{cases}$$

Proof. If

$$\frac{vb}{B_{11}} \equiv \frac{a}{q} \pmod{\mathbb{Z}}$$

with $(a; q) = 1$ then since $(b; B_{11}) = 1$ we must have

$$q = \frac{B_{11}}{(B_{11}; v)} \geqslant \frac{B_{11}}{|v|} \geqslant M_1 N^{1-\delta} D^{-1}.$$

This proves that there are no solutions if $q < N^{1-\delta} M_1 D^{-1}$ so the first part of the lemma follows.

For the remainder of the proof we suppose that $q \geqslant N^{1-\delta} M_1 D^{-1}$. If $(a; q) = 1$ and

$$\frac{vb}{B_{11}} \equiv \frac{a}{q} \pmod{\mathbb{Z}}$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

then $q|B_{11}$. It follows that

$$vb \equiv aB_{11}/q \pmod{B_{11}}.$$

We therefore see that for given q and B_{11} the number of possible v is $O(DN^{-1+\delta}M_1^{-1}+1)$. Moreover, since $q|B_{11}$ there are $O(M_1q^{-1})$ possible B_{11}. By assumption we know that

$$DN^{-1+\delta}M_1^{-1} < 1$$

so we may conclude that the quantity of interest is $O(M_1q^{-1})$ as required. □

Using the last lemma we deduce that the part of ψ_2 with $q \sim Q$, for $Q \geqslant N^{1-\delta}M_1D^{-1}$, is bounded by

$$M_1Q^{-1}\sum_{q\sim Q}\sum_{(a;q)=1}\max_{N'\leqslant N}\left|\sum_{m\leqslant N'}\alpha_m e\left(\frac{am}{q}\right)\right|^2.$$

Applying a maximal form of the large sieve, as given by Montgomery [6], we can majorise this by

$$M_1Q^{-1}N(N+Q^2) = M_1N(Q^{-1}N+Q).$$

Recall that

$$N^{1-\delta}M_1D^{-1} \ll Q \ll M_1$$

so our bound is at most

$$M_1N(N^{\delta}M_1^{-1}D + M_1).$$

We have $M_1 \ll D^{1/2}$ so the first term is always larger and the bound is simply $N^{1+\delta}D$. This holds for all the dyadic intervals $q \sim Q$ under consideration so we conclude that for any $\epsilon > 0$ we have

$$\psi_2(\Lambda, N, \alpha, \delta) \ll_\epsilon N^{1+\delta+\epsilon}D$$

and therefore that

$$\sum_{\lambda\in\Lambda}|\psi_1(\lambda, N, \alpha, \delta)| \ll_\epsilon N^{1+2\delta+\epsilon}M_1^{-1/2}D^{1/2}.$$

This completes the proof of Theorem 2.1.

§3. *Level of distribution.* Rather than considering only the values $f(p, n)$ we will consider values $\alpha_m f(m, n)$ for sequences of complex numbers α_m with $|\alpha_m| \leqslant 1$. Letting α_m be the indicator function of the primes will then recover the case in which we are most interested. Our approach is able to handle any sequence α_m but there are a number of unpleasant technicalities to deal with. To avoid this we will consider only α_m supported on primes m. We will study the quantity

$$A_d(N, \alpha) = \sum_{\substack{(m,n)\in(0,N]\times\mathbb{Z}\\ f(m,n)\equiv 0 \pmod d}}\alpha_m W\left(\frac{n}{N}\right).$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

We expect that for α_m supported on primes we have, at least on average over a suitable range of d,

$$A_d(N,\alpha) \approx M_d(N,\alpha)$$

where

$$M_d(N,\alpha) = \frac{N\nu(d)\hat{W}(0)}{d}\sum_{m\leqslant N}\alpha_m$$

and $\nu(d)$ is the number of solutions, n, of the congruence

$$f(1,n) \equiv 0 \pmod{d}.$$

We therefore wish to estimate the sum

$$\sum_{d\sim D}|A_d(N,\alpha) - M_d(N,\alpha)|.$$

THEOREM 3.1. *Let α_m be a sequence of complex numbers with $|\alpha_m| \leqslant 1$ supported on prime values of m. Suppose $\delta_1 > 0$ and $1 \leqslant D \leqslant N^{4/3-\delta_1}$. There exists a $\delta_2 > 0$ depending only on δ_1 such that*

$$\sum_{d\sim D}|A_d(N,\alpha) - M_d(N,\alpha)| \ll_{\delta_1} N^{2-\delta_2}.$$

The advantage of working with α_m supported on primes is that the contribution to our sum from points (m,n) with $(m;d) > 1$ is small.

LEMMA 3.2. *Under the hypotheses of Theorem 3.1 we have, for any $\epsilon > 0$, that*

$$\sum_{d\sim D}\sum_{\substack{(m,n)\in(0,N]\times\mathbb{Z}\\(m;d)>1,\, f(m,n)\equiv 0 \pmod{d}}}|\alpha_m|W\left(\frac{n}{N}\right) \ll_\epsilon N^{1+\epsilon}.$$

Proof. Since α_m is supported on primes the condition $(m;d) > 1$ implies $m|d$. We therefore have

$$\begin{aligned}
&\sum_{d\sim D}\sum_{\substack{(m,n)\in(0,N]\times\mathbb{Z}\\(m;d)>1,\, f(m,n)\equiv 0 \pmod{d}}}|\alpha_m|W\left(\frac{n}{N}\right)\\
&\ll \sum_{d\sim D}\sum_{\substack{m\leqslant N\\ m|d}}\mathbf{1}_P(m)\sum_{\substack{n\leqslant N\\ f(m,n)\equiv 0 \pmod{d}}}1\\
&= \sum_{m\leqslant N}\mathbf{1}_P(m)\sum_{\substack{n\leqslant N\\ f(m,n)\equiv 0 \pmod{m}}}\#\{d|f(m,n) : d\sim D, m|d\}\\
&\leqslant \sum_{m\leqslant N}\mathbf{1}_P(m)\sum_{\substack{n\leqslant N\\ f(m,n)\equiv 0 \pmod{m}}}\tau(f(m,n))\\
&\ll_\epsilon \sum_{m\leqslant N}\mathbf{1}_P(m)\sum_{\substack{n\leqslant N\\ f(m,n)\equiv 0 \pmod{m}}}N^\epsilon,
\end{aligned}$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

where τ is the divisor function and we have used the fact that f is irreducible so $f(m,n) \neq 0$.

Let f_0 be the coefficient of $n^{\deg f}$ in f. We have

$$\sum_{\substack{m \leqslant N \\ m \mid f_0}} \mathbf{1}_P(m) \sum_{\substack{n \leqslant N \\ f(m,n) \equiv 0 \pmod m}} N^{\epsilon} \ll_f N^{1+\epsilon}.$$

Recall that f is homogeneous. It follows that if a prime m does not divide f_0 but $m \mid f(m,n)$ then we must have $m \mid n$. Therefore

$$\sum_{\substack{m \leqslant N \\ m \nmid f_0}} \mathbf{1}_P(m) \sum_{\substack{n \leqslant N \\ f(m,n) \equiv 0 \pmod m}} N^{\epsilon} = \sum_{\substack{m \leqslant N \\ m \nmid f_0}} \mathbf{1}_P(m) \sum_{\substack{n \leqslant N \\ m \mid n}} N^{\epsilon} \ll_{\epsilon} N^{1+\epsilon}.$$

The result follows. □

Our proof of Theorem 3.1 begins by applying methods from the geometry of numbers, similar to those employed by Daniel in [1]. We call a point (m,n) primitive modulo d if $(m;n;d) = 1$. We say that the primitive points (m_1,n_1) and (m_2,n_2) are equivalent modulo d if

$$(m_2,n_2) \equiv \lambda(m_1,n_1) \pmod d$$

for some $\lambda \in \mathbb{Z}$ which must necessarily satisfy $(\lambda;d) = 1$. We observe that the property $f(m,n) \equiv 0 \pmod d$ is preserved by equivalence so we may let $\mathcal{U}(d)$ be the set of equivalence classes mod d for which it holds.

For each $x \in \mathcal{U}(d)$ we let $\lambda(x)$ be the lattice in $\mathbb{Z}^2$ generated by the points of x. Thus if we fix an $(m,n) \in x$ then $\lambda(x)$ consists of all the points congruent mod d to some multiple of (m,n). It follows that $\det \lambda(x) = d$ and that the set of primitive points in $\lambda(x)$ is precisely x. Each primitive solution of $f(m,n) \equiv 0 \pmod d$ occurs in precisely one lattice $\lambda(x)$ but a nonprimitive solution may occur in more than one. Since any nonprimitive point has $(m;d) > 1$ and $\#\mathcal{U}(d) \ll_{\epsilon,f} d^{\epsilon}$, (see for example Daniel [1, (3.5)]), we can handle this multiplicity issue with the last lemma.

We let $\mathcal{U}'(d)$ be the subset of $\mathcal{U}(d)$ containing those x generated by a point (m,n) with $(m;d) = 1$. If $x \notin \mathcal{U}'(d)$ then all $(m,n) \in \lambda(x)$ have $(m;d) > 1$. It is clear that $\#\mathcal{U}'(d) = \nu(d)$. We can therefore deduce using the last lemma that

$$\sum_{d \sim D} |A_d(N,\alpha) - M_d(N,\alpha)|$$
$$\ll_{\epsilon} N^{1+\epsilon} + \sum_{d \sim D} \sum_{x \in \mathcal{U}'(d)} \left| \sum_{(m,n) \in \lambda(x) \cap (0,N] \times \mathbb{Z}} \alpha_m W\left(\frac{n}{N}\right) - \frac{N \hat{W}(0)}{d} \sum_{m \leqslant N} \alpha_m \right|.$$

We must therefore bound

$$S = \sum_{d \sim D} \sum_{x \in \mathcal{U}'(d)} \left| \psi(\lambda(x), N, \alpha) - \frac{N \hat{W}(0)}{d} \sum_{m \leqslant N} \alpha_m \right|$$

where ψ is the quantity studied in the last section.

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

We let $B_1(x), B_2(x)$ denote the minimal basis of $\lambda(x)$ and write $B(x)$ for the matrix with rows the B_i. If $D \geqslant N^{\delta_1}$ it is necessary to remove from S any lattices for which B_{11} is unusually small, say $B_{11}(x) \leqslant D^{1/2-\eta}$ for some $\eta > 0$. For these lattices we bound the sums

$$\sum_{d\sim D} \sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}} |\psi(\lambda(x), N, \alpha)|$$

and

$$\sum_{d\sim D} \sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}} \left|\frac{N\hat{W}(0)}{d} \sum_{m\leqslant N} \alpha_m\right|.$$

The first sum is bounded by

$$S_1 = \sum_{d\sim D} \sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}} \#(\lambda(x) \cap [0, N]^2)$$

whilst the second is at most of order

$$S_2 = \frac{N^2}{D} \sum_{d\sim D} \sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}} 1.$$

We estimate these using the following lemma.

LEMMA 3.3. *Suppose* $0 \neq (u, v) \in \mathbb{Z}^2$. *Then for any* $\epsilon > 0$ *we have*

$$\#\{(d, x) : d \sim D, x \in \mathcal{U}'(d), (u, v) \in \lambda(x)\} \ll_\epsilon \|(u, v)\|^\epsilon.$$

Proof. Since f is irreducible and $(u, v) \neq 0$ we know that $f(u, v) \neq 0$. The number of possible d is then bounded by

$$\tau(f(u, v)) \ll_\epsilon \|(u, v)\|^\epsilon.$$

For each such d the number of possible x cannot exceed $\nu(d) = O_\epsilon(d^\epsilon)$. The result follows. □

Recall that $\det\lambda(x) \sim D$. Therefore, if $B_{11}(x) \leqslant D^{1/2-\eta}$ we must have $B_1(x) = (u, v)$ for some $0 \neq (u, v) \in \mathbb{Z}^2$ with $u \leqslant D^{1/2-\eta}$ and $v \ll D^{1/2}$. It follows that the number of terms in our sums S_1, S_2 is at most $O_\epsilon(D^{1-\eta}N^\epsilon)$. We immediately deduce that

$$S_2 \ll_\epsilon N^{2+\epsilon} D^{-\eta}.$$

To bound S_1 we use the standard estimate for the number of lattice points to get

$$S_1 \ll \sum_{d\sim D} \sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}} \left(\frac{N^2}{d} + \frac{N}{\|B_1(x)\|} + 1\right).$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

From the above discussion we obtain the bounds

$$\sum_{d\sim D}\sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}}\frac{N^2}{d}\ll_\epsilon N^{2+\epsilon}D^{-\eta}$$

and

$$\sum_{d\sim D}\sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}}1\ll_\epsilon D^{1-\eta}N^\epsilon.$$

Finally we use Lemma 3.3 to get

$$\begin{aligned}\sum_{d\sim D}\sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)\leqslant D^{1/2-\eta}}}&\frac{N}{\|B_1(x)\|}\\ &\ll_\epsilon N^{1+\epsilon}\sum_{0<\|(u,v)\|\ll D^{1/2}}\frac{1}{\sqrt{u^2+v^2}}\\ &\ll_\epsilon N^{1+\epsilon}D^{1/2}.\end{aligned}$$

We conclude that

$$S_1+S_2\ll_\epsilon N^{1+\epsilon}D^{1/2}+D^{1-\eta}N^\epsilon+N^{2+\epsilon}D^{-\eta}$$

so this bound also holds for the contribution to S from lattices with $B_{11}\leqslant D^{1/2-\eta}$. Recalling that $N^{\delta_1}\leqslant D\leqslant N^{4/3-\delta_1}$ we see that if we take a small enough ϵ then this bound is $O(N^{2-\delta_2})$ for $\delta_2>0$ sufficiently small in terms of δ_1 and η. It should be noted that the exponent $\frac{4}{3}$ is not critical for this part of the argument.

It remains to consider

$$S_3=\sum_{d\sim D}\sum_{\substack{x\in\mathcal{U}'(d)\\ B_{11}(x)>D^{1/2-\eta}}}\left|\psi(\lambda(x),N,\alpha)-\frac{N\hat{W}(0)}{d}\sum_{m\leqslant N}\alpha_m\right|,$$

to which we will apply Theorem 2.1. If $D\geqslant N^{\delta_1}$ then η is a quantity that we can take arbitrarily small, whereas if $D\leqslant N^{\delta_1}$ then we shall take $\eta>1/2$ so that all lattices are included.

If $x\in\mathcal{U}'(d)$ then $\lambda(x)$ consists of all points congruent modulo d to a multiple of some (m,n) with $(m;d)=1$. It follows that the m-coordinates of points in $\lambda(x)$ are coprime. The sum is over lattices $\lambda(x)$ which have $\det\lambda(x)\sim D$ and $D^{1/2-\eta}<B_{11}(x)\ll D^{1/2}$. For each possible value of $B_{11}(x)$ in this range there are $O(D^{1/2})$ permissible values for $B_{12}(x)$. It follows by Lemma 3.3 that the number of lattices in the sum with any given value of B_{11} is at most $O_\epsilon(D^{1/2}N^\epsilon)$. We therefore subdivide S_3 into $O_\epsilon(N^\epsilon)$ dyadic intervals depending on the size of B_{11} and then subdivide each dyadic sum into $O_\epsilon(D^{1/2}N^\epsilon)$ subsums in which

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

each possible value of B_{11} occurs at most once. The resulting subsums may be estimated using Theorem 2.1. Suppose $\delta > 0$. If $D \leqslant N^{1-\delta}$ we get

$$S_3 \ll_{\epsilon,A} D^{1/2} N^{\epsilon - A},$$

for any $A \in \mathbb{N}$, which is certainly small enough. If $D \geqslant N^{1-\delta}$ we must check the condition

$$D \leqslant M_1 N^{1-\delta}.$$

However $M_1 \geqslant D^{1/2-\eta}$ so it is sufficient that

$$D \leqslant N^{(1-\delta)/(1/2+\eta)}.$$

Since $D \leqslant N^{4/3-\delta_1}$ this is certainly satisfied if we take δ, η small enough. (Since $D \geqslant N^{1-\delta}$ we are in the case in which any $\eta > 0$ is admissible.) We may therefore deduce from Theorem 2.1 that

$$S_3 \ll_{\epsilon} N^{\epsilon} D^{1/2} \cdot N^{1+2\delta+\epsilon} D^{1/4+\eta/2} \ll_{\epsilon} N^{1+2\delta+\epsilon} D^{3/4+\eta/2}.$$

Since $D \leqslant N^{4/3-\delta_1}$ we see that if we take δ, ϵ and η sufficiently small in terms of δ_1 then

$$S_3 \ll_{\delta_1} N^{2-\delta_2}$$

for some $\delta_2 > 0$. This is where the value $4/3$ is critical as for larger D we do not get a nontrivial bound from Theorem 2.1.

We conclude that

$$\sum_{d \sim D} |A_d(N, \alpha) - M_d(N, \alpha)| \ll_{\delta_1} N^{2-\delta_2}$$

for some $\delta_2 > 0$, thus completing the proof of Theorem 3.1.

When we apply the weighted sieve in the next section we will use the following upper bound to show that not too many values of f are divisible by the square of a prime.

LEMMA 3.4. *Let α_m be a sequence of complex numbers with $|\alpha_m| \leqslant 1$. For any $\delta_1 > 0$ there exists a $\delta_2 > 0$, depending only on δ_1, such that*

$$\sum_{N^{\delta_1} \leqslant p \leqslant N^{2-\delta_1}} |A_{p^2}(N, \alpha)| \ll_{\delta_1} N^{2-\delta_2},$$

the sum being over primes p.

Proof. We have

$$\begin{aligned} A_{p^2}(N, \alpha) &= \sum_{\substack{(m,n) \in (0,N] \times \mathbb{Z} \\ f(m,n) \equiv 0 \pmod{p^2}}} \alpha_m W\left(\frac{n}{N}\right) \\ &\ll \#\{(m, n) \in [0, N]^2 : f(m, n) \equiv 0 \pmod{p^2}\}. \end{aligned}$$

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

If $f(m,n) \equiv 0 \pmod{p^2}$ then $(m,n) \in \lambda(x)$ for at least one $x \in \mathcal{U}(p^2)$. It follows that

$$A_{p^2}(N,\alpha) \ll \sum_{x \in \mathcal{U}(p^2)} \#(\lambda(x) \cap [0,N]^2).$$

We may bound this by

$$\sum_{x \in \mathcal{U}(p^2)} \left(\frac{N^2}{p^2} + \frac{N}{\|B_1(x)\|} + 1 \right).$$

Using that $\#\mathcal{U}(p^2) \ll_\epsilon N^\epsilon$ we have

$$\sum_{N^{\delta_1} \leqslant p \leqslant N^{2-\delta_1}} \sum_{x \in \mathcal{U}(p^2)} \left(\frac{N^2}{p^2} + 1 \right) \ll_\epsilon N^{2-\delta_1+\epsilon}.$$

It therefore remains to estimate

$$N \sum_{N^{\delta_1} \leqslant p \leqslant N^{2-\delta_1}} \sum_{x \in \mathcal{U}(p^2)} \frac{1}{\|B_1(x)\|}.$$

If points are equivalent modulo p^2 then they must also be equivalent modulo p. It follows that if $x \in \mathcal{U}(p^2)$ then there is some $x' \in \mathcal{U}(p)$ with $\lambda(x) \subseteq \lambda(x')$. Different equivalence classes in $\mathcal{U}(p^2)$ may give rise to the same class in $\mathcal{U}(p)$ but the total number of times a class may occur cannot exceed $\#\mathcal{U}(p^2) \ll_\epsilon N^\epsilon$. Our sum is therefore majorised by

$$N^{1+\epsilon} \sum_{N^{\delta_1} \leqslant p \leqslant N^{2-\delta_1}} \sum_{x \in \mathcal{U}(p)} \frac{1}{\|B_1(x)\|}.$$

To estimate this final sum we use part of Daniel's proof of [**1**, Lemma 3.2], which is very similar to our above derivation of a bound on S_1. Specifically, if we set $Q = N^{2-\delta_1}$, our sum is bounded by the quantity $T_1^*(Q)$ defined in that proof so it is $O_\epsilon(N^{1-\delta_1/2+\epsilon})$. We therefore conclude that

$$N^{1+\epsilon} \sum_{N^{\delta_1} \leqslant p \leqslant N^{2-\delta_1}} \sum_{x \in \mathcal{U}(p)} \frac{1}{\|B_1(x)\|} \ll_\epsilon N^{2-\delta_1/2+\epsilon}.$$

The result follows on combining the above estimates and taking $\delta_2 < \delta_1/2$. □

§4. *Proof of Theorem 1.1.* We will sieve the sequence $\mathcal{A} = (a_l)$ given by

$$a_l = \sum_{\substack{(m,n) \in (0,N] \times \mathbb{Z} \\ |f(m,n)| = l}} \mathbf{1}_P(m) W\left(\frac{n}{N}\right).$$

This is supported on $l \ll_f N^k$, and by Theorem 3.1 we know that it has level of distribution N^θ for any $\theta < \frac{4}{3}$. Since f is irreducible we deduce from the prime ideal theorem that the values $\nu(p)$ are 1 on average and we may therefore use

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

a 1-dimensional weighted sieve. By assumption we know that $\nu(p) < p$ for all primes p. It can therefore be shown that

$$\prod_{p<z}\left(1-\frac{\nu(p)}{p}\right)=\frac{c_f+o(1)}{\log z}$$

for some $c_f > 0$.

We use the weighted sieve as described by Greaves in [4, Ch. 5]. If $r \geqslant 2$ we deduce that if

$$\tfrac{3}{4}k < r - \delta_r$$

then for all sufficiently large N we have

$$\sum_l{}^* a_l \gg \frac{N^2}{(\log N)^2},$$

where $\sum^*$ denotes a sum over certain l which are the product of at most r distinct primes. Specifically, [4, §5.2] shows that we can take $\delta_r = 0.144001\ldots$. The above estimate therefore follows if

$$r > \tfrac{3}{4}k + 0.15$$

which is equivalent to $r \geqslant [3k/4] + 1$. Observe that it is essential that we had $\delta_r < \frac{1}{4}$. The simplest form of the weighted sieve [4, §5.1] would therefore have been insufficient.

It remains to show that we can produce numbers with at most r prime factors when counted with multiplicity. Examining the construction of the sieve it can be seen that there are constants $0 < \alpha < \beta < 2$, depending on r, such that $\sum^*$ is actually a sum over l all of whose prime factors exceed N^α and for which

$$\sum_{\substack{p\mid l\\ p\leqslant N^\beta}} 1 + \sum_{p\geqslant N^\beta}\sum_{a:\, p^a\mid l} 1 \leqslant r.$$

This means that only prime factors smaller than N^β are counted without multiplicities. We can deduce from Lemma 3.4 that the contribution of l which are divisible by p^2 for $p \in [N^\alpha, N^\beta]$ is $O(N^{2-\delta})$ for some $\delta > 0$ depending on α and β. We may therefore conclude that for all sufficiently large N we have

$$\sum_{l\in P_r} a_l \gg \frac{N^2}{(\log N)^2}$$

thereby completing the proof of Theorem 1.1.

Acknowledgements. This work was completed as part of my DPhil, for which I was funded by EPSRC grant EP/P505666/1. I am very grateful to the EPSRC for funding me and to my supervisor, Roger Heath-Brown, for all his valuable help and advice.

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press

References

1. S. Daniel, On the divisor-sum problem for binary forms. *J. Reine Angew. Math.* **507** (1999), 107–129.
2. E. Fouvry and H. Iwaniec, Gaussian primes. *Acta Arith.* **79**(3) (1997), 249–287.
3. G. Greaves, Large prime factors of binary forms. *J. Number Theory* **3** (1971), 35–59.
4. G. Greaves, *Sieves in Number Theory* (*Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **43**), Springer (Berlin, 2001).
5. D. R. Heath-Brown, Primes represented by $x^3 + 2y^3$. *Acta Math.* **186**(1) (2001), 1–84.
6. H. L. Montgomery, Maximal variants of the large sieve. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **28**(3) (1982), 805–812 (1981).
7. H.-E. Richert, Selberg's sieve with weights. *Mathematika* **16** (1969), 1–22.

A. J. Irving,
Mathematical Institute,
University of Oxford,
Andrew Wiles Building,
Radcliffe Observatory Quarter,
Woodstock Road,
Oxford OX2 6GG,
U.K.
E-mail: irving@maths.ox.ac.uk

https://doi.org/10.1112/S002557931400031X Published online by Cambridge University Press