Illinois Journal of Mathematics
Volume 59, Number 4, Winter 2015, Pages 999–1023
S 0019-2082

THUE EQUATIONS AND LATTICES

JEFFREY LIN THUNDER

ABSTRACT. We consider Diophantine equations of the kind $|F(x,y)| = m$, where $F(X,Y) \in \mathbb{Z}[X,Y]$ is a homogeneous polynomial of degree at least 3 that has non-zero discriminant, m is a fixed positive integer and x, y are relatively prime integer solutions. Our results improve upon previous theorems due to Bombieri and Schmidt and also Stewart. We further provide reasonable heuristics for conjectures of Schmidt and Stewart regarding such equations.

1. Introduction

Suppose $F(X,Y) \in \mathbb{Z}[X,Y]$ is a homogeneous polynomial of degree $d \geq 3$ that has non-zero discriminant and m is a positive integer. In this paper we are concerned with the number of primitive, that is, x and y are relatively prime, solutions $(x,y) \in \mathbb{Z}^2$ to the Thue equation

$$|F(x,y)| = m. \tag{1}$$

Thue in [9] famously showed that the number of such solutions is necessarily finite under the hypothesis that F is irreducible over $\mathbb{Q}$. In fact, his method enabled one to derive an upper bound on the number of solutions, an upper bound that would depend on m and the polynomial F. Indeed, Lewis and Mahler in [4] provided just such a bound. Their bound was an explicit function of m, d and the height of F. Previous to the result of Lewis and Mahler, Siegel had made the conjecture that an upper bound could be obtained that was independent of the particular coefficients of the polynomial F. Evertse proved this conjecture in his doctoral thesis (see [3]). A few years later Bombieri and Schmidt [2] improved markedly on Evertse's bound, showing that the number of primitive solutions to (1) is no more than some fixed (absolute) constant multiple of $d^{1+\omega(m)}$, where $\omega(m)$ denotes the number of distinct prime factors

Received March 15, 2016; received in final form November 4, 2016.
2010 *Mathematics Subject Classification.* 11D75, 11J25.

of m. Given that typically $\omega(m)$ is roughly $\log\log m$ for large m, Schmidt posited (see [7, Chapter 3, conjecture]) that the number of primitive solutions to (1) should be bounded above by some multiple (possibly depending on F) of a power of $\log m$ when $m > 1$.

Two years after the publication of Bombieri and Schmidt's result, Stewart [8] provided a bound that was often (depending on the prime factorization of the parameter m) much stronger than the bound of Bombieri and Schmidt. Stewart's main result involved a somewhat complicated quantity, but one can easily state the following consequence. In what follows, $D(F)$ denotes the discriminant of F.

THEOREM (Stewart). *Suppose $F(X,Y) \in \mathbb{Z}[X,Y]$ is a homogeneous polynomial of degree $d \geq 3$ with non-zero discriminant and content* 1. *Let $\varepsilon > 0$. Suppose m is a positive integer and m' is a divisor of m relatively prime to $D(F)$ that satisfies $(m')^{1+\varepsilon} \geq m^{(2/d)+\varepsilon}/|D(F)|^{1/d(d-1)}$. Then the number of primitive solutions to* (1) *is at most*

$$\left(5600d + \frac{1400}{\varepsilon}\right) d^{\omega(m')}.$$

The constants 5600 and 1400 here carry no particular importance beyond specificity. The major improvement over the result of Bombieri and Schmidt is that the quantity $\omega(m')$ is possibly much smaller than $\omega(m)$. See [5] for related improvements. In the same paper, Stewart explicitly constructed forms of various degrees to demonstrate lower bounds for the number of primitive solutions to (1). In so doing he was lead to the following conjecture (see Section 6 of [8]): there is an absolute constant c_1 such that, for all forms F as in the theorem above, there is a positive c_2 (depending on F) such that (1) has at most c_1 primitive solutions for all $m \geq c_2$.

We will prove that primitive solutions to (1) occur in certain sublattices of determinant almost m (the "almost" due to possible common prime factors of m and the discriminant $D(F)$). Moreover, for F of degree at least 4 and m sufficiently large we will show that the existence of a primitive solution to (1) in such a sublattice implies that either the primitive solution gives rise to an exceptional approximation in the sense of Thue–Siegel–Roth or the sublattice has an exceptionally small first successive minima in the sense of Minkowski. In order to state our main results, we introduce a bit of notation.

Denote the set of places of $\mathbb{Q}$ by $M(\mathbb{Q})$. For any $v \in M(\mathbb{Q})$ we let $|\cdot|_v$ denote the usual v-adic absolute value on $\mathbb{Q}$ and $\mathbb{Q}_v$ denote the topological completion of $\mathbb{Q}$ with respect to this absolute value, though we will continue to use $|\cdot|$ for the usual Euclidean absolute value. We fix algebraic closures $\overline{\mathbb{Q}}_v$ for each of these and assume that our original v-adic absolute value on $\mathbb{Q}$ is extended to $\overline{\mathbb{Q}}_v$ in the usual manner. We follow the standard convention of identifying the finite places with positive primes.

Any form (i.e., homogeneous polynomial) $F(X,Y) \in \mathbb{Q}[X,Y]$ factors completely into a product of linear forms over some splitting field:

$$F(X,Y) = \prod_{i=1}^{d} L_i(X,Y).$$

This splitting field may be embedded into any $\overline{\mathbb{Q}}_v$; we abuse notation somewhat and write the above for the factorization of F over $\overline{\mathbb{Q}}_v$ for all places $v \in M(\mathbb{Q})$. These linear factors are only unique up to a scalar multiple, of course. As usual, we say a linear form $L_i(X,Y)$ is defined over $\mathbb{Q}_v$ if all possible quotients of coefficients are in $\mathbb{Q}_v$. For any form $F(X,Y) \in \mathbb{Z}[X,Y]$ and place $v \in M(\mathbb{Q})$, set $c_F(v)$ to be the number of linear factors that are defined over $\mathbb{Q}_v$. For any integer $m > 1$, set

$$c_F(m) = \prod_{\substack{p \mid m \\ p \text{ prime}}} c_F(p).$$

Finally, given a form F with non-zero discriminant and positive integer m, we will often deal with two particular divisors of m determined by the prime factors common to m and $D(F)$. For notational convenience, we will set m_F to be the largest divisor of m such that $|m_F|_p < |D(F)|_p$ for all primes $p \mid m_F$ (assuming $m \nmid D(F)$ so that such a divisor exists) and

$$D_F(m) = \prod_{\substack{p \text{ prime} \\ |m|_p < |D(F)|_p}} \frac{|D(F)|_p^{1/2}}{|m|_p}.$$

We have $D_F(m_F) = D_F(m)$ directly from the definitions.

THEOREM 1. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree $d \geq 2$ with non-zero discriminant and content 1 and suppose m is a positive integer with $|m|_p < |D(F)|_p$ for all primes $p \mid m$. Then the primitive $\mathbf{x} \in \mathbb{Z}^2$ with $m \mid F(\mathbf{x})$ are contained in $c_F(m)$ sublattices of $\mathbb{Z}^2$ of determinant $D_F(m)$. In particular, for all positive $m \nmid D(F)$ the primitive solutions to (1) lie in $c_F(m_F)$ sublattices of $\mathbb{Z}^2$ of determinant $D_F(m)$.*

We note here that there are no solutions to (1) if $c_F(m_F) = 0$. In other words, $c_F(m_F) = 0$ implies that there is some local obstruction to solving (1).

One way we may utilize Theorem 1 to bound the number of primitive solutions to (1) is if there is a sufficiently large divisor m' of m (akin to that in Stewart's result) with $c_F(m')$ small. In fact, given a sublattice $\Lambda \subseteq \mathbb{Z}^2$ of relatively large determinant, we can provide an upper bound not just on the number of primitive solutions to (1), but even to the related inequality

$$\left|F(x,y)\right| \leq m. \tag{1'}$$

THEOREM 2. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree $d \geq 3$ with non-zero discriminant and content* 1. *Suppose m is a positive integer and $\Lambda \subseteq \mathbb{Z}^2$ is a sublattice with* $\det(\Lambda) = Am^{2/d}/|D(F)|^{1/d(d-1)}$ *for some $A > 0$. If $A \geq 5^4$, then the number of primitive points in Λ that are solutions to* (1′) *is less than*

$$2 + 2d\left(13 + \frac{31}{\log(d-1)} + \frac{\log(\frac{2\log m}{d\log A} + 2)}{\log(d-1)}\right).$$

If $A < 5^4$, then the number of solutions is less than

$$\frac{2\cdot 5^4}{A}\left(2 + 2d\left(13 + \frac{31}{\log(d-1)} + \frac{\log(\frac{\log m}{2d\log 5} + 2)}{\log(d-1)}\right)\right).$$

COROLLARY. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree $d \geq 3$ with non-zero discriminant and content* 1. *Suppose m is a positive integer and m' is a divisor of m relatively prime to $D(F)$ that satisfies $m' = Am^{2/d}/|D(F)|^{1/d(d-1)}$ for some $A \geq 1$. Then the number of primitive solutions $(x,y) \in \mathbb{Z}^2$ to* (1′) *with $m'|F(x,y)$ is less than*

$$2500d\left(59 + \frac{\log(2 + \log m/(1 + \log A))}{\log(d-1)}\right)c_F(m').$$

Proof. One readily checks that for $d \geq 3$, $(d/2)\log A \geq 1 + \log A$ if $A \geq 5^4$ and $2d\log 5 \geq 1 + \log A$ if $1 \leq A < 5^4$. The rest follows from Theorems 1 and 2. □

Comparing the corollary with the result of Stewart above, our constants are in the same ballpark even though we are estimating more than just the solutions to (1). Moreover, we can easily replace the 2500 with 2 once $A \geq 5^4$. Our major novelty is replacing Stewart's $d^{\omega(m')}$ term with the explicit $c_F(m')$ and the implicit information on the underlying lattice(s). Clearly $c_F(m') \leq d^{\omega(m')}$ always, but will typically be much smaller, though we remark that a thorough reading of Stewart's proof shows that one could use $c_F(m')$ there (indeed, the proof in [2] shows that their $d^{1+\omega(m)}$ term may be replaced with $dc_F(m)$). Also, for large m the estimate in Stewart's result tends to a multiple (depending on F) of $d^{\omega(m')}\log m$ as the divisor m' approaches $m^{2/d}/|D(F)|^{1/d(d-1)}$, whereas our estimate is bounded above by a multiple of $c_F(m')\log\log m$. Again, the most important observation here regards the lattices. It is this that allows us to provide concrete support for the conjectures of Stewart and Schmidt above.

We will prove Theorems 1 and 2 in the next two sections. In Section 4 we will use Theorem 1 together with gap arguments to provide reasonable heuristics for Stewart's conjecture above when the degree $d \geq 5$, and for a strong form of Schmidt's conjecture when $d = 4$.

2. Lattices arising from Thue equations

Our goal in this section is to prove Theorem 1. For the remainder of this paper, we will use a bold-faced font to denote the coefficient vector of a form (e.g., $\mathbf{F}$ for the coefficient vector of a form F). We let $\|\cdot\|_p$ denote the supremum norm for vectors over $\overline{\mathbb{Q}}_p$ for primes p and let $\|\cdot\|$ denote the usual Euclidean norm for vectors over $\mathbb{C}$. Given a form $F(X,Y)$ with its factorization into linear forms, $F(X,Y)=\prod_{i=1}^d L_i(X,Y)$, we set

$$\mathcal{H}(F)=\prod_{i=1}^d \|\mathbf{L}_i\|.$$

We note that

$$\|\mathbf{F}\|_p=\prod_{i=1}^d \|\mathbf{L}_i\|_p$$

for all primes p by Gauss' lemma.

Our first step in the proof of Theorem 1 is the following non-archimedean version of [10, Lemma 4]. In what follows, for any vector $\mathbf{v}$ we denote the transpose (column vector) by $\mathbf{v}^{\mathrm{tr}}$.

LEMMA 1. *Let K be a topologically complete field with respect to a non-archimedean absolute value $|\cdot|'$ and $L_1(\mathbf{X}),\dots,L_n(\mathbf{X})\in K[\mathbf{X}]$ be n linearly independent linear forms in n variables. Let $\|\cdot\|'$ denote the supremum norm on K^n. Suppose $\mathbf{x}\in K^n$ and j is such that*

$$\frac{|L_j(\mathbf{x})|'}{\|\mathbf{L}_j\|'}\ge\frac{|L_i(\mathbf{x})|'}{\|\mathbf{L}_i\|'}$$

for $i=1,\dots,n$. Then

$$\frac{|L_j(\mathbf{x})|'}{\|\mathbf{L}_j\|'}\ge\frac{\|\mathbf{x}\|'|\det(\mathbf{L}_1^{\mathrm{tr}},\dots,\mathbf{L}_n^{\mathrm{tr}})|'}{\prod_{i=1}^n\|\mathbf{L}_i\|'}.$$

Proof. The statement is obvious if $\mathbf{x}=\mathbf{0}$, so suppose otherwise. Then without loss of generality, we may assume $\|\mathbf{L}_i\|'=1$ for all i and $\|\mathbf{x}\|'=1$. Let T denote the $n\times n$ matrix with rows $\mathbf{L}_i$ and write

$$\mathfrak{m}=\min_{\substack{\mathbf{y}\in K^n\\ \|\mathbf{y}\|'=1}}\left\{\left\|T\mathbf{y}^{\mathrm{tr}}\right\|'\right\},\qquad \mathfrak{M}=\max_{\substack{\mathbf{y}\in K^n\\ \|\mathbf{y}\|'=1}}\left\{\left\|T\mathbf{y}^{\mathrm{tr}}\right\|'\right\}.$$

Suppose $\|T\mathbf{x}_1^{\mathrm{tr}}\|'=\mathfrak{m}$ and $\|\mathbf{x}_1\|'=1$. Choose $\mathbf{x}_2,\dots,\mathbf{x}_n\in K^n$, all of length 1, that also satisfy $|\det(\mathbf{x}_1^{\mathrm{tr}},\dots,\mathbf{x}_n^{\mathrm{tr}})|'=1$. We then have

$$\begin{aligned}\left|\det(T)\right|'&=\left|\det(T)\right|'\left|\det\left(\mathbf{x}_1^{\mathrm{tr}},\dots,\mathbf{x}_n^{\mathrm{tr}}\right)\right|'=\left|\det\left(T\mathbf{x}_1^{\mathrm{tr}},\dots,T\mathbf{x}_n^{\mathrm{tr}}\right)\right|'\\ &\le\prod_{l=1}^n\left\|T\mathbf{x}_l^{\mathrm{tr}}\right\|'\\ &\le\mathfrak{m}\mathfrak{M}^{n-1}.\end{aligned}$$

Since $\|\mathbf{L}_i\|' = 1$ for all i and the absolute value is non-archimedean we have $\mathfrak{M} \le 1$, so that $\mathfrak{m} \ge |\det(T)|'$. On the other hand, by our choice of j we also have $|L_j(\mathbf{x})|' \ge |L_i(\mathbf{x})|'$ for all $i = 1, \ldots, n$. Since $\|\cdot\|'$ is the supremum norm, these n inequalities (and the definition of T) imply that $|L_j(\mathbf{x})|' \ge \|T\mathbf{x}^{\mathrm{tr}}\|' \ge \mathfrak{m}$. Thus

$$\left|L_j(\mathbf{x})\right|' \ge \mathfrak{m} \ge \left|\det(T)\right|' = \left|\det\left(\mathbf{L}_1^{\mathrm{tr}}, \ldots, \mathbf{L}_n^{\mathrm{tr}}\right)\right|'. \qquad \square$$

LEMMA 2. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree $d \ge 2$ with non-zero discriminant and write*

$$F(X,Y) = \prod_{i=1}^{d} L_i(X,Y),$$

where $L_i(X,Y)$ is a linear form for all $i = 1, \ldots, d$. Suppose $\mathbf{x} \in \mathbb{Q}^2$. Then for any place $v \in M(\mathbb{Q})$, if i_0 is an index with

$$\frac{|L_{i_0}(\mathbf{x})|_v}{\|\mathbf{L}_{i_0}\|_v} = \min_{1 \le i \le d} \left\{ \frac{|L_i(\mathbf{x})|_v}{\|\mathbf{L}_i\|_v} \right\},$$

then

$$\frac{|L_{i_0}(\mathbf{x})|_v}{\|\mathbf{L}_{i_0}\|_v} \le \begin{cases} \frac{2^{d-1}|F(\mathbf{x})|\mathcal{H}(F)^{d-2}}{\|\mathbf{x}\|^{d-1}|D(F)|^{1/2}} & \text{if } v = \infty, \\ \frac{|F(\mathbf{x})|_v \|\mathbf{F}\|_v^{d-2}}{\|\mathbf{x}\|_v^{d-1}|D(F)|_v^{1/2}} & \text{otherwise.} \end{cases}$$

Further, if $v = p$ is a prime with

$$\frac{|F(\mathbf{x})|_p}{\|\mathbf{F}\|_p} < \frac{\|\mathbf{x}\|_p^d |D(F)|_p}{\|\mathbf{F}\|_p^{2(d-1)}},$$

then the index i_0 above is unique and L_{i_0} is defined over $\mathbb{Q}_p$.

Proof. By Lemma 1 if $v \ne \infty$ and [10, Lemma 4] if $v = \infty$, we have

$$\frac{|L_i(x,y)|_v}{\|\mathbf{L}_i\|_v} \ge \begin{cases} \frac{\|\mathbf{x}\| |\det(\mathbf{L}_{i_0}^{\mathrm{tr}}, \mathbf{L}_i^{\mathrm{tr}})|}{2\|\mathbf{L}_{i_0}\|\|\mathbf{L}_i\|} & \text{if } v = \infty, \\ \frac{\|\mathbf{x}\|_v |\det(\mathbf{L}_{i_0}^{\mathrm{tr}}, \mathbf{L}_i^{\mathrm{tr}})|_v}{\|\mathbf{L}_{i_0}\|_v \|\mathbf{L}_i\|_v} & \text{otherwise} \end{cases} \tag{2}$$

for all $i \ne i_0$, since the linear forms $L_1(X,Y), \ldots, L_d(X,Y)$ are pairwise linearly independent (because the discriminant is non-zero). We also have Hadamard's inequality

$$\frac{|\det(\mathbf{L}_i^{\mathrm{tr}}, \mathbf{L}_j^{\mathrm{tr}})|_v}{\|\mathbf{L}_i\|_v \|\mathbf{L}_j\|_v} \le 1$$

for all places $v \in M(\mathbb{Q})$ and all indices i, j.

If $v=\infty$, then by (2), Hadamard's inequality and the definition of $\mathcal{H}(F)$ we get

$$\begin{aligned}\frac{|F(\mathbf{x})|}{\mathcal{H}(F)} &= \frac{|F(\mathbf{x})|}{\|\mathbf{L}_1\|\cdots\|\mathbf{L}_d\|} = \prod_{i=1}^{d}\frac{|L_i(\mathbf{x})|}{\|\mathbf{L}_i\|}\\ &\geq \frac{|L_{i_0}(\mathbf{x})|}{\|\mathbf{L}_{i_0}\|}\prod_{i\neq i_0}\frac{\|\mathbf{x}\||\det(\mathbf{L}_{i_0}^{\mathrm{tr}},\mathbf{L}_i^{\mathrm{tr}})|}{2\|\mathbf{L}_{i_0}\|\|\mathbf{L}_i\|}\\ &= \frac{\|\mathbf{x}\|^{d-1}|L_{i_0}(\mathbf{x})|}{2^{d-1}\|\mathbf{L}_{i_0}\|}\prod_{i\neq i_0}\frac{|\det(\mathbf{L}_{i_0}^{\mathrm{tr}},\mathbf{L}_i^{\mathrm{tr}})|}{\|\mathbf{L}_{i_0}\|\|\mathbf{L}_i\|}\\ &\geq \frac{\|\mathbf{x}\|^{d-1}|L_{i_0}(\mathbf{x})|}{2^{d-1}\|\mathbf{L}_{i_0}\|}\prod_{i>j}\frac{|\det(\mathbf{L}_{i}^{\mathrm{tr}},\mathbf{L}_j^{\mathrm{tr}})|}{\|\mathbf{L}_{i}\|\|\mathbf{L}_j\|}\\ &= \frac{\|\mathbf{x}\|^{d-1}|L_{i_0}(\mathbf{x})|}{2^{d-1}\|\mathbf{L}_{i_0}\|}\frac{|D(F)|^{1/2}}{\mathcal{H}(F)^{d-1}}.\end{aligned}$$

If $v\neq\infty$ then in a similar manner but using Gauss' lemma with $\|\mathbf{F}\|_v$ in place of $\mathcal{H}(F)$ above, we see that

$$\frac{|F(\mathbf{x})|_v}{\|\mathbf{F}\|_v} \geq \frac{\|\mathbf{x}\|_v^{d-1}|L_{i_0}(\mathbf{x})|_v}{\|\mathbf{L}_{i_0}\|_v}\frac{|D(F)|_v^{1/2}}{\|\mathbf{F}\|_v^{d-1}}.$$

This proves the first part of the lemma.

Now suppose $v=p$ a prime and there is an index $i_1\neq i_0$ with

$$\frac{|L_{i_1}(\mathbf{x})|_p}{\|\mathbf{L}_{i_1}\|_p} = \min_{1\leq i\leq d}\left\{\frac{|L_i(\mathbf{x})|_p}{\|\mathbf{L}_i\|_p}\right\}.$$

Then Lemma 1 would also give

$$\frac{|L_{i_0}(\mathbf{x})|_p}{\|\mathbf{L}_{i_0}\|_p} \geq \frac{\|\mathbf{x}\|_p|\det(\mathbf{L}_{i_0}^{\mathrm{tr}},\mathbf{L}_{i_1}^{\mathrm{tr}})|_p}{\|\mathbf{L}_{i_0}\|_p\|\mathbf{L}_{i_1}\|_p}. \tag{3}$$

Using (2), (3) and Hadamard's inequality yields

$$\begin{aligned}\frac{|F(\mathbf{x})|_p}{\|\mathbf{F}\|_p} &= \frac{|F(\mathbf{x})|_p}{\|\mathbf{L}_1\|_p\cdots\|\mathbf{L}_d\|_p} = \prod_{i=1}^{d}\frac{|L_i(\mathbf{x})|_p}{\|\mathbf{L}_i\|_p}\\ &\geq \|\mathbf{x}\|_p^d\frac{|\det(\mathbf{L}_{i_1}^{\mathrm{tr}},\mathbf{L}_{i_0}^{\mathrm{tr}})|_p}{\|\mathbf{L}_{i_0}\|_p\|\mathbf{L}_{i_1}\|_p}\prod_{i\neq i_0}\frac{|\det(\mathbf{L}_{i_0}^{\mathrm{tr}},\mathbf{L}_i^{\mathrm{tr}})|_p}{\|\mathbf{L}_{i_0}\|_p\|\mathbf{L}_i\|_p}\\ &\geq \|\mathbf{x}\|_p^d\prod_{i\neq j}\frac{|\det(\mathbf{L}_{i}^{\mathrm{tr}},\mathbf{L}_j^{\mathrm{tr}})|_p}{\|\mathbf{L}_{i}\|_p\|\mathbf{L}_j\|_p}\\ &= \frac{\|\mathbf{x}\|_p^d|D(F)|_p}{\|\mathbf{L}_1\|_p^{2(d-1)}\cdots\|\mathbf{L}_d\|_p^{2(d-1)}}\\ &= \frac{\|\mathbf{x}\|_p^d|D(F)|_p}{\|\mathbf{F}\|_p^{2(d-1)}}.\end{aligned}$$

Therefore if

$$\frac{\|F(\mathbf{x})\|_p}{\|\mathbf{F}\|_p} < \frac{\|\mathbf{x}\|_p^d |D(F)|_p}{\|\mathbf{F}\|_p^{2(d-1)}}$$

the index i_0 is unique. Finally, to see that L_{i_0} is defined over $\mathbb{Q}_p$, suppose this were not the case. Then without loss of generality (i.e., possibly introducing a scalar multiple) there would be a σ in the Galois group of $\overline{\mathbb{Q}}_p$ over $\mathbb{Q}_p$ with $\sigma(L_{i_0}) = L_{i_1}$ for some $i_1 \neq i_0$ between 1 and d. We then have $|L_{i_1}(\mathbf{x})|_p = |L_{i_0}(\mathbf{x})|_p$ and $\|\mathbf{L}_{i_1}\|_p = \|\mathbf{L}_{i_0}\|_p$ (see [1, Chapter 2, Theorem 7]), which contradicts what we have already proven. $\square$

Lemma 3. *Suppose p is a prime and $L(X,Y) \in \mathbb{Z}_p[X,Y]$ is a linear form with $\|\mathbf{L}\|_p = 1$. Then for all integers $c \geq 0$ the set*

$$S_p := \left\{\mathbf{z} \in \mathbb{Z}_p^2 : \left|L(\mathbf{z})\right|_p \leq p^{-c}\right\}$$

is a $\mathbb{Z}_p$-submodule with index $[\mathbb{Z}_p^2 : S_p] = p^c$.

Proof. Suppose $\mathbf{z}_1, \mathbf{z}_2 \in S_p$ and $z \in \mathbb{Z}_p$. Then

$$\left|L(\mathbf{z}_1 + \mathbf{z}_2)\right|_p = \left|L(\mathbf{z}_1) + L(\mathbf{z}_2)\right|_p \leq \max\left\{\left|L(\mathbf{z}_1)\right|_p, \left|L(\mathbf{z}_2)\right|_p\right\} \leq p^{-c}$$

and

$$\left|L(z\mathbf{z}_1)\right|_p = \left|zL(\mathbf{z}_1)\right|_p = |z|_p \left|L(\mathbf{z}_1)\right|_p \leq p^{-c},$$

so that S_p is a $\mathbb{Z}_p$-module.

Write $L(X,Y) = a_1 X + a_2 Y$. Then $\|\mathbf{L}\|_p = \max\{|a_1|_p, |a_2|_p\} = 1$. Suppose first that $|a_1|_p = 1$ and set $\mathbf{z}_1 = (1,0)$ and $\mathbf{z}_2 = (a_2, -a_1)$. Then $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}_p^2$ and so $\mathbb{Z}_p^2 \supseteq \{x\mathbf{z}_1 + y\mathbf{z}_2 : x, y \in \mathbb{Z}_p\}$. On the other hand, given $(b_1, b_2) \in \mathbb{Z}_p^2$ we can write

$$(b_1, b_2) = \left(b_1 + b_2 a_2 a_1^{-1}\right)\mathbf{z}_1 - b_2 a_1^{-1}\mathbf{z}_2.$$

Since $|a_1|_p = 1$, $a_1^{-1} \in \mathbb{Z}_p$ and so $b_1 + b_2 a_2 a_1^{-1}, b_2 a_1^{-1} \in \mathbb{Z}_p$. We thus see that we have the reverse inclusion as well: $\mathbb{Z}_p^2 \subseteq \{x\mathbf{z}_1 + y\mathbf{z}_2 : x, y \in \mathbb{Z}_p\}$. Hence, $\mathbb{Z}_p^2 = \{x\mathbf{z}_1 + y\mathbf{z}_2 : x, y \in \mathbb{Z}_p\}$. Now using $L(\mathbf{z}_2) = 0$ and $|L(\mathbf{z}_1)|_p = |a_1|_p = 1$, we see that

$$S_p = \left\{x\mathbf{z}_1 + y\mathbf{z}_2 : x, y \in \mathbb{Z}_p, |x|_p \leq p^{-c}\right\}.$$

This shows that the index $[\mathbb{Z}_p^2 : S_p]$ is the index of $\{x \in \mathbb{Z}_p : |x|_p \leq p^{-c}\}$ in $\mathbb{Z}_p$, which is p^c.

If $|a_1|_p < 1$, then we must have $|a_2|_p = 1$ and in this case we set $\mathbf{z}_1 = (0,1)$. The proof in this case is exactly as above but with $a_2^{-1} \in \mathbb{Z}_p$ now. $\square$

Lemma 4. *Let $\mathcal{P}$ be a finite set of prime numbers and for every $p \in \mathcal{P}$ let $L_p(X,Y) \in \mathbb{Z}_p[X,Y]$ be a linear form with $\|\mathbf{L}_p\|_p = 1$. For each $p \in \mathcal{P}$ let a_p be a non-negative integer and set*

$$S_p = \left\{(x,y) \in \mathbb{Z}_p^2 : \left|L_p(x,y)\right|_p \leq p^{-a_p}\right\}.$$

Set $S_p = \mathbb{Z}_p^2$ *for all primes* $p \notin \mathcal{P}$. *Then*

$$\Lambda := \bigcap_{p \text{ prime}} \mathbb{Q}^2 \cap S_p$$

is a sublattice of $\mathbb{Z}^2$ *with*

$$\det(\Lambda) = \prod_{p \in \mathcal{P}} p^{a_p}.$$

Proof. Since $S_p = \mathbb{Z}_p^2$ for all but finitely many primes p, by [12, Chapter V, Theorem 2] Λ is the unique sublattice of $\mathbb{Z}^2$ whose closure in $\mathbb{Q}_p^2$ for all primes p is S_p. (Using the notation of that result, set $k = \mathbb{Q}$, $E = \mathbb{Q}^2$, $L = \mathbb{Z}^2$, $M = \Lambda$, $L_p = \mathbb{Z}_p^2$ and $M_p = S_p$ for all primes p.) We thus may compute the index of Λ in $\mathbb{Z}^2$ via [12, Chapter V, Theorem 4, Corollary 1]:

$$\left[\mathbb{Z}^2 : \Lambda\right] = \prod_{p \text{ prime}} \left[\mathbb{Z}_p^2 : S_p\right].$$

Now $\det(\Lambda) = [\mathbb{Z}^2 : \Lambda]$ by definition and $[\mathbb{Z}_p^2 : S_p] = p^{a_p}$ if $p \in \mathcal{P}$ by Lemma 3. If $p \notin \mathcal{P}$ then $[\mathbb{Z}_p^2 : S_p] = [\mathbb{Z}_p^2 : \mathbb{Z}_p^2] = 1$ by construction. The lemma follows. □

Proof of Theorem 1. If $\mathbf{x} \in \mathbb{Z}^2$ with $m | F(\mathbf{x})$ and p is a prime dividing m, then $|F(\mathbf{x})|_p \le |m|_p < |D(F)|_p$. If $\mathbf{x}$ is a primitive point and the content of F is 1, then by Lemma 2 there is a unique linear factor $L_p(X,Y) \in \mathbb{Q}_p[X,Y]$ of F with

$$\frac{|L_p(\mathbf{x})|_p}{\|\mathbf{L}_p\|_p} \le \frac{|F(\mathbf{x})|_p}{|D(F)|_p^{1/2}} \le \frac{|m|_p}{|D(F)|_p^{1/2}}$$

since now $\|\mathbf{x}\|_p = 1 = \|\mathbf{F}\|_p$. After possibly rescaling may assume without loss of generality that our linear factor $L_p(X,Y) \in \mathbb{Z}_p[X,Y]$, $\|\mathbf{L}_p\|_p = 1$, and whence

$$\big|L_p(\mathbf{x})\big|_p \le \frac{|F(\mathbf{x})|_p}{|D(F)|_p^{1/2}} \le \frac{|m|_p}{|D(F)|_p^{1/2}}.$$

We do this for all primes p dividing m and then invoke Lemma 4, obtaining a sublattice Λ of $\mathbb{Z}^2$ that contains our primitive point $\mathbf{x}$ and has determinant

$$\det(\Lambda) = \prod_{\substack{p \text{ prime} \\ p | m}} \frac{|D(F)|_p^{1/2}}{|m|_p} = D_F(m).$$

On the other hand, there are $c_F(p)$ possible linear factors $L_p(X,Y)$ here for each prime p by definition, whence $c_F(m)$ possible sublattices when we consider all primes dividing m. □

3. Proof of Theorem 2

If $F(X,Y)$ is any form and $\Lambda = \mathbb{Z}\mathbf{z}_1 \oplus \mathbb{Z}\mathbf{z}_2$ is a lattice, then considering solutions $\mathbf{z} \in \Lambda$ to (1′) is the same as considering solutions $(x,y) \in \mathbb{Z}^2$ to the inequality $|F_\Lambda(x,y)| \le m$, where the form $F_\Lambda(X,Y) := F(X\mathbf{z}_1 + Y\mathbf{z}_2)$. The choice of basis is not unique here of course. We may also view $F_\Lambda(X,Y)$ as a composition $F \circ T$, where $T \in \mathrm{GL}_2(\mathbb{R})$ sends the canonical basis of $\mathbb{Z}^2$ to a basis of Λ. Note that a different choice of basis amounts to multiplying T by an element of $\mathrm{GL}_2(\mathbb{Z})$.

In addition to $\mathcal{H}(F)$, our proof will involve two additional heights; we define

$$\mathcal{M}(F) = \min_{T \in \mathrm{GL}_2(\mathbb{Z})} \mathcal{H}(F \circ T) \quad \text{and}$$
$$\mathfrak{m}(F) = \min_{\substack{T \in \mathrm{GL}_2(\mathbb{R}) \\ |\det(T)|=1}} \mathcal{H}(F \circ T).$$

We remark that in general (see [11, Lemma 1]) for any form F of degree d and any $T \in \mathrm{GL}_2(\mathbb{R})$,

$$\begin{aligned} D(F \circ T) &= D(F)\det(T)^{d(d-1)}, \\ \mathfrak{m}(F \circ T) &= \mathfrak{m}(F)\big|\det(T)\big|^{d/2}, \\ \mathcal{M}(F) \ge \mathfrak{m}(F) &\ge \big|D(F)\big|^{1/2(d-1)}. \end{aligned} \tag{4}$$

In particular, we see that $|D(F_\Lambda)|$, $\mathfrak{m}(F_\Lambda)$ and $\mathcal{M}(F_\Lambda)$ are all well-defined (i.e., are independent of the particular choice of basis). For a given positive integer m we set $\mathcal{M}(F_\Lambda, m)$ to be the minimum of $\mathcal{H}(F_\Lambda)$ over all bases $\mathbf{z}_1, \mathbf{z}_2$ of Λ with $\mathbf{z}_1$ a solution to (1′), assuming such a primitive solution exists.

The main idea for determining solutions to (1′) is to say that some linear factor of F must be relatively small for a given solution. For example, suppose we rewrite $F(X,Y) = a\prod_{i=1}^{d}(X - \alpha_i Y)$. Now if $(x,y) \in \mathbb{Z}^2$ is any solution to (1′) with $y \ne 0$, then

$$|\alpha_i - x/y| \le \frac{d2^{d-1}m\mathcal{H}(F)^{d-2}}{|y|^d|D(F)|^{1/2}} = d2^{d-1}\big(\mathcal{H}(F)/m\big)^{d-2}\frac{m^{d-1}}{|D(F)|^{1/2}}\frac{1}{|y|^d} \tag{5}$$

for some index i by [6, Chapter 3, Lemmas 3A and 3B]. Considering the inequality (5), one can see that the major goal is to estimate those solutions $\mathbf{x} = (x,y)$ to (1′) with $|y|$ small, so that any remaining solutions may be dealt with using gap arguments and ultimately a quantitative version of Roth's theorem. We will use the following as the main part of our proof of Theorem 2.

PROPOSITION. *Suppose* $F(X,Y) \in \mathbb{Z}[X,Y]$ *is a form of degree* $d \ge 3$ *with non-zero discriminant and content* 1, m *is a positive integer and* $\Lambda \subseteq \mathbb{Z}^2$ *is a lattice with* $\det(\Lambda) = Am^{2/d}/|D(F)|^{1/d(d-1)}$. *If* Λ *contains a primitive solution to* (1′), *then* $\mathcal{M}(F_\Lambda, m) \ge A^{d/2}m$ *and if* $A \ge 5^4$ *the number of primitive* *solutions to* (1′) *in* Λ *is less than*

$$2+2d\biggl(13+\frac{\log 2^{10}3^3 5^3}{\log(d-1)}+\frac{\log 2^9 3^3 5^2}{\log(d-5/4)} + \frac{\log\bigl(\frac{\log m}{\log(\mathcal{M}(F_\Lambda,m))-\log m}+2\bigr)}{\log(d-1)}\biggr).$$

The proof of the proposition will rely on a few lemmas, though we note that the inequalities $\mathcal{M}(F_\Lambda,m)\geq\mathcal{M}(F_\Lambda)\geq A^{d/2}m$ follow directly from the definitions, (4) and the hypotheses. We assume $\Lambda=\mathbb{Z}\mathbf{z}_0\oplus\mathbb{Z}\mathbf{z}_0'$ where $\mathbf{z}_0$ is a primitive solution to (1′) and $\mathcal{M}(F_\Lambda,m)=\mathcal{H}(F_\Lambda)$. We will write

$$F(X,Y)=\prod_{i=1}^{d}L_i(X,Y),$$
$$F_\Lambda(X,Y)=\prod_{i=1}^{d}\bigl(XL_i(\mathbf{z}_0)+YL_i\bigl(\mathbf{z}_0'\bigr)\bigr)=F(\mathbf{z}_0)\prod_{i=1}^{d}(X+\alpha_iY),$$

where $\alpha_i=L_i(\mathbf{z}_0')/L_i(\mathbf{z}_0)$. For notational convenience, in what follows we will denote the quantity $\mathcal{M}(F_\Lambda,m)/m$ by B. The hypothesis that $A\geq 5^4$ thus implies that $B\geq 5^{2d}$.

With the above conventions in place, we see by (4) and (5) that for any solution $\mathbf{z}=x\mathbf{z}_0+y\mathbf{z}_0'\in\Lambda$ to (1′) with $y\neq 0$ there is some index i with

$$\begin{aligned}|\alpha_i-x/y| &\leq\frac{d2^{d-1}m^{d-1}B^{d-2}}{|y|^d|D(F_\Lambda)|^{1/2}}\\ &=\frac{d2^{d-1}m^{d-1}B^{d-2}}{|y|^d|D(F)|^{1/2}\det(\Lambda)^{d(d-1)/2}}\\ &\leq\frac{d2^{d-1}B^{d-2}}{|y|^d(5^4)^{d(d-1)/2}}\\ &<\frac{B^{d-2}}{2|y|^d}.\end{aligned}\tag{6}$$

We may utilize a standard gap principle argument to estimate those solutions with $|y|>B$ (see Lemma 7 below). Eventually we come to the point where a quantitative version of Roth's theorem is invoked (Lemma 8). But before we do that, we deal with those solutions where $|y|$ is smaller. The following is a variation on [6, Chapter 3, Lemma 5B].

LEMMA 5. *For every primitive lattice point* $\mathbf{z}=x\mathbf{z}_0+y\mathbf{z}_0'\in\Lambda$ *with* $y\neq 0$ *that is a solution to* (1′), *there are* $\psi_1(\mathbf{z}),\dots,\psi_d(\mathbf{z})\in[0,1]$ *that, if not zero, are at least* $1/(2d)$, *satisfy* $\sum_{i=1}^{d}\psi_i(\mathbf{z})\geq 1/2$, *and also*

$$\frac{|L_i(\mathbf{z}_0)|}{|L_i(\mathbf{z})|}\geq\bigl(B^{\psi_i(\mathbf{z})}-2\bigr)|y|$$

for all $i=1,\dots,d$.

Proof. We first claim that $2|L_{i_0}(\mathbf{z}_0)| \le |L_{i_0}(\mathbf{z})|$ for some index i_0. Indeed, if this were not the case then $\Lambda' := \mathbb{Z}\mathbf{z}_0 \oplus \mathbb{Z}\mathbf{z}$ is a sublattice of Λ and $F_{\Lambda'}(X,Y) := F(X\mathbf{z}_0 + Y\mathbf{z})$ satisfies

$$\begin{aligned}\mathcal{H}(F_{\Lambda'})^2 &= \prod_{i=1}^{d} \bigl|L_i(\mathbf{z}_0)\bigr|^2 + \bigl|L_i(\mathbf{z})\bigr|^2 \\ &< \prod_{i=1}^{d} 5\bigl|L_i(\mathbf{z}_0)\bigr|^2 \le 5^d m^2\end{aligned}$$

since $\mathbf{z}_0$ is a solution to (1′). But now by (4) and the hypotheses we have a contradiction:

$$\begin{aligned}5^{d/2} m > \mathcal{H}(F_{\Lambda'}) &\ge \mathfrak{m}\bigl(F'\bigr) \\ &= \mathfrak{m}(F)\det\bigl(\Lambda'\bigr)^{d/2} \\ &\ge \mathfrak{m}(F)\det(\Lambda)^{d/2} \\ &\ge \bigl|D(F)\bigr|^{1/2(d-1)}\det(\Lambda)^{d/2} \\ &\ge 5^{2d} m.\end{aligned}$$

With the claim shown, choose an index i_0 with $2|L_{i_0}(\mathbf{z}_0)| \le |L_{i_0}(\mathbf{z})|$. Since $\mathbf{z}$ is a primitive lattice point there is a $\mathbf{z}' \in \Lambda$ with $\Lambda = \mathbb{Z}\mathbf{z} \oplus \mathbb{Z}\mathbf{z}'$. Further, we may add any integer multiple of $\mathbf{z}$ to $\mathbf{z}'$ here. Thus, we may choose $\mathbf{z}'$ such that $\alpha := \Re(L_{i_0}(\mathbf{z}')/L_{i_0}(\mathbf{z}))$ satisfies $|\alpha| \le 1/2$. We now write $\mathbf{z}_0 = z\mathbf{z} + z'\mathbf{z}'$ for some $z, z' \in \mathbb{Z}$ with $|z'| = [\Lambda : \mathbb{Z}\mathbf{z}_0 \oplus \mathbb{Z}\mathbf{z}] = |y|$. For any linear form $L(X,Y)$ we have

$$\frac{L(\mathbf{z}_0)}{L(\mathbf{z})} = z + z'\frac{L(\mathbf{z}')}{L(\mathbf{z})}.$$

In particular, using $L = L_{i_0}$ we see that $|z + z'\alpha| \le 1/2$, and for all $i = 1, \ldots, d$

$$\begin{aligned}\frac{|L_i(\mathbf{z}_0)|}{|L_i(\mathbf{z})|} &= \left|z + z'\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| \\ &= \left|z'\left(\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})} - \alpha\right) + z + z'\alpha\right| \\ &\ge \left|z'\left(\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})} - \alpha\right)\right| - |z + z'\alpha| \\ &\ge |z'|\left(\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| - \frac{1}{2}\right) - \frac{1}{2} \\ &\ge |y|\left(\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| + 1 - 2\right).\end{aligned} \tag{7}$$

Since $|F(\mathbf{z})| \le m$ and $\Lambda = \mathbb{Z}\mathbf{z} \oplus \mathbb{Z}\mathbf{z}'$,

$$\begin{aligned}
\prod_{i=1}^{d}\left(1+\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right|\right) &\ge \prod_{i=1}^{d}\sqrt{1+\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right|^2} \\
&= \frac{1}{|F(\mathbf{z})|}\prod_{i=1}^{d}\sqrt{\left|L_i(\mathbf{z})\right|^2+\left|L_i(\mathbf{z}')\right|^2} \\
&\ge \frac{\mathcal{M}(F_\Lambda, m)}{|F(\mathbf{z})|} \\
&\ge B.
\end{aligned} \tag{8}$$

We define $\psi_i(\mathbf{z})$ by

$$B^{\psi_i(\mathbf{z})} = \begin{cases} B & \text{if } \left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| + 1 \ge B, \\ 1 & \text{if } \left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| + 1 < B^{1/(2d)}, \\ \left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| + 1 & \text{otherwise.} \end{cases}$$

Now by construction $0 \le \psi_i(\mathbf{z}) \le 1$ for all $i = 1, \dots, d$ and any $\psi_j(\mathbf{z}) \ge 1/2d$ if it isn't zero. We have $\sum_{i=1}^{d} \psi_i(\mathbf{z}) \ge 1$ if any $\psi_j(\mathbf{z}) = 1$, so suppose $\psi_i(\mathbf{z}) < 1$ for all $i = 1, \dots, d$. Then by (8)

$$\begin{aligned}
B^{1/2}\prod_{i=1}^{d} B^{\psi_i(\mathbf{z})} &> \prod_{\substack{1\le i\le d\\ \psi_i(\mathbf{z})=0}}\left(1+\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right|\right)\prod_{\substack{1\le i\le d\\ \psi_i(\mathbf{z})>0}} B^{\psi_i(\mathbf{z})} \\
&= \prod_{1\le i\le d}\left(1+\left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right|\right) \\
&\ge B.
\end{aligned}$$

This shows that $\sum_{i=1}^{d} \psi_i(\mathbf{z}) \ge 1/2$ in all cases. Also by construction $B^{\psi_i(\mathbf{z})} \le \left|\frac{L_i(\mathbf{z}')}{L_i(\mathbf{z})}\right| + 1$ for all i, so that the remaining desired inequalities follow from (7). $\square$

LEMMA 6. *For all $c > 0$ there are less than $2d(2c+1)$ primitive solutions $\mathbf{z} = x\mathbf{z}_0 + y\mathbf{z}_0' \in \Lambda$ to (1′) with $y \neq 0$ and $|y| \le B^c$.*

We thus are able to rather efficiently estimate solutions where $|y| \le B^c$ for any fixed constant c. In particular, though it's certainly possible to improve upon particular aspects of Lemma 5, there wouldn't be much to gain (the exception being if one could improve upon B, specifically, if one could replace B by a larger quantity in terms of m or F). However, we remark that the hypothesis $\det(\Lambda) \ge 5^4 m^{2/d}|D(F)|^{1/d(d-1)}$ can be relaxed to $\det(\Lambda) \ge 5^4 (m/\mathfrak{m}(F))^{2/d}$, both in Lemma 5 and here in Lemma 6.

Proof. By Lemma 5

$$\frac{|L_i(\mathbf{z})|}{|L_i(\mathbf{z}_0)|} = |\alpha_i - x/y| \le \frac{1}{(B^{\psi_i(\mathbf{z})}-2)|y|^2} \tag{9}$$

for all solutions $\mathbf{z} = x\mathbf{z}_0 + y\mathbf{z}_0' \in \Lambda$ to (1′) with $y \neq 0$ and all $i = 1, \ldots, d$.

Let $\mathcal{S}$ denote the set of primitive solutions $\mathbf{z} = x\mathbf{z}_0 + x\mathbf{z}_0' \in \Lambda$ to (1′) with $1 \le y \le B^c$. For the moment fix an index i and consider the sum $\sum \psi_i(\mathbf{z})$ over all $\mathbf{z} \in \mathcal{S}$. Obviously we may restrict to solutions with $\psi_i(\mathbf{z}) \neq 0$; we arrange these solutions $\mathbf{z}_l = x_l\mathbf{z}_0 + y_l\mathbf{z}_0', l = 1, \ldots, n$ so that $y_l \le y_{l+1}$ for all l. Then by Lemma 5 and (9)

$$\begin{aligned}
\frac{1}{|y_l y_{l+1}|} &\le \left|\frac{x_l}{y_l} - \frac{x_{l+1}}{y_{l+1}}\right| \\
&\le \left|\alpha_i - \frac{x_l}{y_l}\right| + \left|\alpha_i - \frac{x_{l+1}}{y_{l+1}}\right| \\
&\le \frac{1}{(B^{\psi_i(\mathbf{z}_l)}-2)|y_l|^2} + \frac{1}{(B^{\psi_i(\mathbf{z}_{l+1})}-2)|y_{l+1}|^2} \\
&\le \frac{1}{(B^{\psi_i(\mathbf{z}_l)}-2)|y_l|^2} + \frac{1}{(B^{\psi_i(\mathbf{z}_{l+1})}-2)|y_l y_{l+1}|},
\end{aligned}$$

whence

$$|y_{l+1}| \ge \big(B^{\psi_i(\mathbf{z}_l)}-2\big)\Big(1 - \big(B^{\psi_i(\mathbf{z}_{l+1})}-2\big)^{-1}\Big)|y_l|. \tag{10}$$

Since $\psi_i(\mathbf{z}_l) \ge 1/2d$ for all our $\mathbf{z}_l$ and $B \ge 5^{2d}$, we have $B^{\psi_i(\mathbf{z}_l)} \ge 5$ and thus $B^{\psi_i(\mathbf{z}_l)} - 3 \ge B^{\psi_i(\mathbf{z}_l)\log 2/\log 5}$. We now repeatedly apply (10) to get

$$\begin{aligned}
B^c &\ge |y_n| \\
&\ge \big(B^{\psi_i(\mathbf{z}_1)}-2\big)\big(B^{\psi_i(\mathbf{z}_2)}-3\big)\cdots\big(B^{\psi_i(\mathbf{z}_{n-1})}-3\big)\Big(1-\big(B^{\psi_i(\mathbf{z}_n)}-2\big)^{-1}\Big)|y_1| \\
&> \left(\prod_{l=1}^{n-1}\big(B^{\psi_i(\mathbf{z}_l)}-3\big)\right) \times \big(1-(1/3)\big) \\
&\ge (2/3)\prod_{l=1}^{n-1} B^{\psi_i(\mathbf{z}_l)\log 2/\log 5}.
\end{aligned}$$

Taking logarithms yields

$$c + \log_B(3/2) > \sum_{l=1}^{n-1} \psi_i(\mathbf{z}_l)\log 2/\log 5,$$

and since $\psi_i(\mathbf{z}_n) \le 1$,

$$c + \log_B(3/2) + \log 2/\log 5 > \sum_{l=1}^{n} \psi_i(\mathbf{z}_l)\log 2/\log 5 = \sum_{\mathbf{z}\in\mathcal{S}} \psi_i(\mathbf{z})\log 2/\log 5.$$

Finally, by Lemma 5 and this last inequality

$$\begin{aligned}
|\mathcal{S}| &\le \sum_{\mathbf{z}\in\mathcal{S}}\sum_{i=1}^{d} 2\psi_i(\mathbf{z}) \\
&= \sum_{i=1}^{d}\sum_{\mathbf{z}\in\mathcal{S}} 2\psi_i(\mathbf{z}) \\
&< \sum_{i=1}^{d} 2c + 2\log_B(3/2) + 2\log 2/\log 5 \\
&\le \sum_{i=1}^{d} 2c + 2\log(3/2)/\log\bigl(5^{2d}\bigr) + 2\log 2/\log 5 \\
&= \sum_{i=1}^{d} 2c + \log(3/2)/d\log 5 + 2\log 2/\log 5 \\
&\le \sum_{i=1}^{d} 2c + \log(3/2)/3\log 5 + 2\log 2/\log 5 \\
&< d(2c+1).
\end{aligned}$$

The same argument works for estimating the number of primitive solutions $x\mathbf{z}_0 + y\mathbf{z}_0'$ to $(1')$ with $1 \le -y \le B^c$. □

LEMMA 7. *For all $C_2 > C_1 > B$, the number of primitive solutions $\mathbf{z} = x\mathbf{z}_0 + y\mathbf{z}_0' \in \Lambda$ to $(1')$ with $C_1 \le |y| \le C_2$ is less than*

$$2d\left(1 + \frac{\log(\log C_2/\log(C_1/B))}{\log(d-1)}\right).$$

Proof. We will use (6). Suppose $x\mathbf{z}_0 + y\mathbf{z}_0', x'\mathbf{z}_0 + y'\mathbf{z}_0' \in \Lambda$ are primitive solutions to $(1')$ with both

$$|\alpha_i - x/y| < \frac{B^{d-2}}{2|y|^d}, \qquad |\alpha_i - x'/y'| < \frac{B^{d-2}}{2|y'|^d}$$

for some index i. Suppose further that $y' \ge y > 0$. Then by the inequalities above

$$\begin{aligned}
\frac{1}{|yy'|} &\le \left|\frac{x}{y} - \frac{x'}{y'}\right| \\
&\le \left|\alpha_i - \frac{x}{y}\right| + \left|\alpha_i - \frac{x'}{y'}\right| \\
&< \frac{B^{d-2}}{2|y|^d} + \frac{B^{d-2}}{2|y'|^d} \le \frac{B^{d-2}}{|y|^d},
\end{aligned}$$

so that $|y'| \geq |y|^{d-1}/B^{d-2}$. Hence, if $x_1\mathbf{z}_0 + y_1\mathbf{z}'_0, x_2\mathbf{z}_0 + y_2\mathbf{z}'_0, \ldots$ are primitive solutions to (1′) as above with $C_1 \leq y_1 \leq y_2 \leq \cdots \leq C_2$, then repeatedly applying the above inequality yields

$$C_2 \geq y_{l+1} \geq \frac{y_1^{(d-1)^l}}{B^{((d-1)^{l-1}+\cdots+1)(d-2)}} \geq \frac{C_1^{(d-1)^l}}{B^{(d-1)^l-1}} > (C_1/B)^{(d-1)^l}$$

for all $l \geq 1$. We take logarithms twice to get

$$\frac{\log(\log C_2/(\log C_1/B))}{\log(d-1)} > l.$$

Taking into account the d possible indices i and employing the same argument for solutions with $y < 0$ gives the lemma. □

Lemma 8. *There are fewer than*

$$2d\left(6 + \frac{\log 2^9 3^3 5^2}{\log(d-5/4)}\right)$$

primitive $\mathbf{z} = x\mathbf{z}_0 + y\mathbf{z}'_0 \in \Lambda$ *solutions to* (1′) *with*

$$|y| \geq \max\bigl\{B^{4(d-1)}, \bigl(8^d \mathcal{M}(F_\Lambda, m)\bigr)^{2^{10}3^3 5^3}\bigr\}.$$

Proof. We note that the α_i are conjugate algebraic numbers with absolute height $h(\alpha_i)$ satisfying

$$h(\alpha_i)^d = \mathcal{H}(F_\Lambda) = \mathcal{M}(F_\Lambda, m)$$

(see [6, Chapter 3, Lemma 2A], for example). Given a solution as in the lemma, using the hypothesis $|y| \geq B^{4(d-1)}$ and (6) yields

$$\begin{aligned} |\alpha_i - x/y| &\leq \frac{d2^{d-1}B^{d-2}}{|y|^d} \\ &< \frac{B^{d-1}}{2|y|^d} \leq \frac{1}{2|y|^{d-1/4}} \end{aligned} \tag{11}$$

for some index i. We claim that

$$|\alpha_i - x/y| < \bigl(h(x/y)\bigr)^{-\sqrt{2d}(1+1/20)}, \tag{12}$$

where $h(x/y) = \sqrt{x^2+y^2}$ is the (absolute) height of x/y. To see this, we first note that $|x/y| < |\alpha_i| + 1$, so that $h(x/y) < (|\alpha_i| + 2)|y| \leq 3h(\alpha_i)^d|y|$. Since $d \geq 3$, one readily verifies that $d - 1/4 \geq \sqrt{2d}(1 + 1/10)$. Using this we easily get $(3h(\alpha_i)^d)^{d-1/4} < (3h(\alpha_i)^d)^{\sqrt{d}} < y^{\sqrt{2d}/20}$ (with quite a bit of room to spare, in fact). In addition, we also get

$$\begin{aligned} y^{d-1/4} &\geq \left(\frac{h(x/y)}{3h(\alpha_i)^d}\right)^{d-1/4} \\ &> h(x/y)^{\sqrt{2d}(1+1/10)} y^{-\sqrt{2d}/20} \end{aligned}$$

$$\begin{aligned}&\geq h(x/y)^{\sqrt{2d}(1+1/10)}h(x/y)^{-\sqrt{2d}/20}\\&=h(x/y)^{\sqrt{2d}(1+1/20)}.\end{aligned}$$

Therefore, (12) follows from (11).

According to [6, Chapter 2, Theorem 6] (with $m=2$ and $\chi=1/20$ there), the rational solutions x/y to (12) satisfy $h(x/y)\leq(8h(\alpha_i))^{d2^{10}3^35^3}$ or $w\leq h(x/y)<w^{2^93^35^2d^2}$ for some $w>1$. The first option here is ruled out for us by hypothesis since $h(x/y)\geq|y|$. Hence, it remains to estimate the number of primitive solutions (x,y) to (11) with $w/(3h(\alpha_i)^d)\leq|y|<w^{2^93^35^2d^2}$. We clearly may assume that $w\geq(8h(\alpha_i))^{4d}$.

Suppose $(x_0,y_0),(x_1,y_1),\dots$ are the primitive solutions to (11) with $y_i>0$ and arranged so that $0<y_0\leq y_1\leq\cdots$. We then have

$$\begin{aligned}\frac{1}{|y_ly_{l+1}|}&\leq\left|\frac{x_l}{y_l}-\frac{x_{l+1}}{y_{l+1}}\right|\\&\leq\left|\alpha_i-\frac{x_l}{y_l}\right|+\left|\alpha_i-\frac{x_{l+1}}{y_{l+1}}\right|\\&<\frac{1}{2|y_l|^{d-1/4}}+\frac{1}{2|y_{l+1}|^{d-1/4}}\\&\leq\frac{1}{|y_l|^{d-1/4}},\end{aligned}$$

so that $|y_{l+1}|\geq|y_l|^{d-5/4}$ for all $l\geq0$. Moreover, since $w\geq(8^dh(\alpha_i)^d)^4$ and $d\geq3$ we have

$$w^{d-2}\geq\bigl(8h(\alpha_i)\bigr)^{4d}>\bigl(3h(\alpha_i)^d\bigr)^{2(d-2)}B^{2(d-2)}\geq\bigl(3h(\alpha_i)^dB\bigr)^{d-1},$$

so that also by (11)

$$y_1\geq\frac{y_0^{d-1}}{B^{d-1}}\geq\left(\frac{w}{3h(\alpha_i)^dB}\right)^{d-1}>\frac{w^{d-1}}{w^{d-2}}=w.$$

We thus have $y_l\geq w^{(d-5/4)^{l-1}}$ for all $l\geq1$. Now since all $y_l<w^{2^93^35^2d^2}$ and $d\geq3$, we must have

$$l<1+\frac{\log 2^93^35^2d^2}{\log(d-5/4)}<5+\frac{\log 2^93^35^2}{\log(d-5/4)}.$$

Considering the d possible indices i above and accounting for those solutions with $y<0$ in the same manner completes the proof. □

Proof of the proposition. We first set $c=2$ in Lemma 6 to see that the number of primitive solutions $\mathbf{z}=x\mathbf{z}_0+y\mathbf{z}_0'\in\Lambda$ to ($1'$) with $1\leq|y|\leq B^2$ is less than $10d$. Next we set $C_1=B^2$ and $C_2=B^{4(d-1)}$ in Lemma 7 to see that the number of solutions with $B^2 \le |y| \le B^{4(d-1)}$ is less than

$$2d\left(1+\frac{\log(\log B^{4(d-1)}/\log B)}{\log(d-1)}\right) = 2d\big(2+\log 4/\log(d-1)\big).$$

If on the other hand, we set $C_2 = (8^d \mathcal{M}(F_\Lambda, m))^{2^{10}3^3 5^3}$, then (recall $B \ge 5^{2d} > 8^d$) the number of solutions with $B^2 \le |y| \le (8^d \mathcal{M}(F_\Lambda, m))^{2^{10}3^3 5^3}$ is less than

$$\begin{aligned}
&2d\left(1+\frac{\log(2^{10}3^3 5^3 \log(8^d \mathcal{M}(F_\Lambda, m))/\log B)}{\log(d-1)}\right)\\
&\quad< 2d\left(1+\frac{\log(2^{10}3^3 5^3 (1+\log(\mathcal{M}(F_\Lambda, m))/\log B))}{\log(d-1)}\right)\\
&\quad= 2d\left(1+\frac{\log(2^{10}3^3 5^3 (2+\log m/\log B))}{\log(d-1)}\right)\\
&\quad= 2d\left(1+\frac{\log 2^{10}3^3 5^3}{\log(d-1)}+\frac{\log(2+\log m/\log B)}{\log(d-1)}\right).
\end{aligned}$$

Therefore the number of solutions with

$$B^2 \le |y| \le \max\left\{B^{4(d-1)}, \big(8^d \mathcal{M}(F_\Lambda, m)\big)^{2^{10}3^3 5^3}\right\}$$

is less than

$$2d\left(2+\frac{\log 2^{10}3^3 5^3}{\log(d-1)}+\frac{\log(2+\log m/\log B)}{\log(d-1)}\right).$$

Combining this with Lemma 8, the number of solutions with $y \ne 0$ is less than

$$10d + 2d\left(8+\frac{\log 2^{10}3^3 5^3}{\log(d-1)}+\frac{\log 2^9 3^3 5^2}{\log(d-5/4)}+\frac{\log(2+\log m/\log B)}{\log(d-1)}\right).$$

Of course we also have the two solutions $\pm\mathbf{z}_0$ as well, giving the proposition. $\square$

Proof of Theorem 2. Suppose first that $A \ge 5^4$. We may assume that there is a primitive solution $(x, y) \in \Lambda$ to (1′). We apply the proposition, noting that

$$\log\big(\mathcal{M}(F_\Lambda, m)/m\big) \ge \log\big(A^{d/2}\big), \qquad \frac{\log(2^{10}3^3 5^3)}{\log(d-1)}+\frac{\log(2^9 3^3 5^2)}{\log(d-5/4)} < \frac{31}{\log(d-1)}$$

since $d \ge 3$. For $A < 5^4$ we use the proposition in conjuction with Lemma 2C (and Remark 2D) of [6, Chapter 3] as follows. Let p be any prime satisfying $(5^4/A) \le p \le 2(5^4/A) - 1$ and let F be a form as in the proposition except that $A < 5^4$. Then there are $p+1$ forms G with $|D(G)| = |D(F)|p^{d(d-1)}$ and any primitive integer solution (x, y) to (1′) is a primitive integral solution to $|G(x, y)| \le m$ for one of these forms G. Since

$$\det(\Lambda) = \frac{Am^{2/d}}{|D(F)|^{1/d(d-1)}} = \frac{Am^{2/d}p}{|D(G)|^{1/d(d-1)}} \ge \frac{5^4 m^{2/d}}{|D(G)|^{1/d(d-1)}},$$

we may apply the proposition to these $p+1 \le 2(5^4/A)$ forms G to prove the case of Theorem 2 when $A < 5^4$. □

4. Heuristics for the conjectures of Stewart and Schmidt

When considering the conjectures of Stewart and Schmidt, it is convenient to segregate off the solutions that arise from exceptionally good Diophantine approximations. Specifically, write the form F as a product of linear forms as before:

$$F(X,Y) = \prod_{i=1}^{d} L_i(X,Y).$$

For $\varepsilon > 0$, we say a non-zero $\mathbf{x} \in \mathbb{Z}^2$ is ε-*exceptional* if

$$\frac{|L_i(\mathbf{x})|}{\|\mathbf{L}_i\|} < \frac{1}{\|\mathbf{x}\|^{1+\varepsilon}}$$

for some index i. In other words the ε-exceptional points are the points dealt with by Roth's theorem. It is well known that the number of such exceptional points is bounded above by an explicit function of ε and F (see [6, Chapter 2, Theorem 9B], for example), thus justifying the "exceptional" moniker.

According to Lemma 2 in Section 2 above, any solution $\mathbf{x}$ to ($1'$) must be ε-exceptional once $\|\mathbf{x}\|$ is (up to constants depending on F) about $m^{1/(d-2-\varepsilon)}$. If $d \ge 5$, this implies via Theorem 1 that primitive solutions to (1) which are not ε-exceptional are lattice points of relatively small length. At the expense of an added $\log\log m$ factor, we can deal with solutions to ($1'$) with length down to about $m^{1/(d-2)}$ (see Lemma 9 below), and at the further expense of an arbitrarily small power of $\log m$, we can even deal with solutions to (1) with length down to about $m^{1/(d-2)}/(\log m)^{\delta}$ for a given fixed positive δ. This is explicitly carried out in the case $d = 4$ in Theorem 4 below. For forms of degree 4 this implies, via Theorem 1 once more, that with no more than $(\log m)^{\varepsilon}$ exceptions our primitive solutions to (1) are once again lattice points of relatively small length.

In what follows we will make the above notions precise and then show (Theorem 5) that this gives good evidence for Stewart's conjecture when the degree $d \ge 5$ and a strong form of Schmidt's conjecture when the degree $d = 4$. The proofs for the following results will be given after the statement of Theorem 5 so as to not interupt the flow of thought.

THEOREM 3. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous polynomial of degree $d \ge 5$ with non-zero discriminant and content* 1, *m be a positive integer and* $0 < \varepsilon < d-4$. *Suppose* $0 < 2\delta < d-4-\varepsilon$, *$m'|m$ is a divisor of m satisfying*

$$\left(2D_F(m')/\pi\right)^{1+\delta} \ge \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}$$

and $\Lambda\subseteq\mathbb{Z}^2$ *is a lattice with* $\det(\Lambda)=D_F(m')$. *If there is a primitive solution* $\mathbf{x}\in\Lambda$ *to* (1′) *that is not* ε*-exceptional, then* $\lambda_2>\lambda_1$ *and*

$$\|\mathbf{x}\|=\lambda_1\le\big(2D_F\big(m'\big)/\pi\big)^{(1+\delta)/(d-2-\varepsilon)},$$

where $\lambda_1\le\lambda_2$ *are the successive minima of* Λ *with respect to the unit disk.*

Corollary. *Let* $F(X,Y)\in\mathbb{Z}[X,Y]$ *be a homogeneous polynomial of degree* $d\ge 5$ *with non-zero discriminant and content* 1, m *be a positive integer not dividing* $D(F)$ *and* $0<\varepsilon<d-4$. *If*

$$2D_F(m)/\pi>\left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{2/(d-2-\varepsilon)}$$

then the primitive solutions to (1) *that are not* ε*-exceptional lie in* $c_F(m_F)$ *sublattices of determinant* $D_F(m)$. *Moreover, given one of these sublattices* Λ *there is at most one pair* $\pm\mathbf{x}\in\Lambda$ *of primitive solutions to* (1) *that are not* ε*-exceptional, and this can only be the case if the first minima* λ_1 *of* Λ *satisfies*

$$\|\mathbf{x}\|=\lambda_1\le\left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{1/(d-2-\varepsilon)}.$$

Theorem 4. *Let* $F(X,Y)\in\mathbb{Z}[X,Y]$ *be a homogeneous polynomial of degree* 4 *with non-zero discriminant and content* 1, m *be a positive integer,* $0<\varepsilon<1$ *and* $0<\delta$. *If* $m'\nmid D(F)$ *is any divisor of* m *satisfying* $D_F(m')>(\log m)^{3\delta}$, *then the number of primitive solutions* $\mathbf{x}$ *to* (1) *that satisfy*

$$\|\mathbf{x}\|\ge\frac{e^{1/3}4m^{1/2}\mathcal{H}(F)}{|D(F)|^{1/4}(\log m)^{\delta}}$$

and are not ε*-exceptional is less than*

$$8\left(1+c_F\big(m'_F\big)+\frac{\log\log(2^{3/2}m^{1/2}\mathcal{H}(F)/|D(F)|^{1/4})}{\log 3}\right).$$

If further

$$2D_F(m)/\pi>\frac{e^{2/3}16m\mathcal{H}(F)^2}{|D(F)|^{1/2}(\log m)^{2\delta}},$$

then any remaining primitive solutions to (1) *that are not* ε*-exceptional lie in* $c_F(m_F)$ *sublattices* Λ *of determinant* $D_F(m)$. *Given one of these sublattices* Λ *there is at most one pair* $\pm\mathbf{x}\in\Lambda$ *of such primitive solutions and this can only be the case if the first minima* λ_1 *of* Λ *satisfies* $\|\mathbf{x}\|=\lambda_1$.

Take a form F as in Theorem 2. Note that for any given $\varepsilon>0$, if m is sufficiently large we have no ε-exceptional primitive solutions to (1). Suppose first that the degree $d\ge 5$. Choose a positive $\gamma<(1/2)-1/(d-2)$, set

$\delta = (1/2) - 1/(d-2) - \gamma$ and let ε be given by $\frac{1}{d-2-\varepsilon} = \frac{1}{d-2} + \frac{\delta}{2}$. For m sufficiently large (depending on F and γ) we have $m \nmid D(F)$ and

$$\begin{aligned}\left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{1/(d-2-\varepsilon)} &= \left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{1/(d-2)+\delta/2} \\ &\le \big(2D_F(m)/\pi\big)^{1/(d-2)+\delta} \\ &= \big(2D_F(m)/\pi\big)^{1/2-\gamma}\end{aligned}$$

and no ε-exceptional primitive solutions to (1). Now by the corollary to Theorem 3 and Theorem 1 , all primitive solutions to (1) are in certain sublattices Λ of determinant $D_F(m)$, each such sublattice can contain at most one pair of primitive solutions $\pm\mathbf{x}$, and this only if the first minima of the sublattice satisfies $\|\mathbf{x}\| = \lambda_1 \le (\det(\Lambda))^{(1/2)-\gamma}$.

Now suppose the degree of the form F is 4 and choose $\varepsilon, \gamma > 0$. For m sufficiently large (depending on F) there are no 1/2-exceptional primitive solutions to (1). For m sufficiently large (depending on F and γ), any divisor m' of m with $m' > (\log m)^{7\gamma}$ satisfies $m' \nmid D(F)$ and $D_F(m') > (\log m)^{6\gamma}$. For m sufficiently large (depending on ε and γ), the least divisor m' of m greater than $(\log m)^{7\gamma}$ satisfies $4^{\omega(m')} \le 16(\log m)^{\varepsilon}$, so that $c_F(m') \le 16(\log m)^{\varepsilon}$. For m sufficiently large (depending on F and ε)

$$\frac{\log\log(2^{3/2}m^{1/2}\mathcal{H}(F)/|D(F)|^{1/4})}{\log 3} \le (\log m)^{\varepsilon}.$$

Finally, for m sufficiently large (depending on F and γ)

$$\begin{aligned}\big(2D_F(m)/\pi\big)^{1/2} &> \frac{D_F(m)^{1/2}}{(\log D_F(m))^{\gamma}} \\ &\ge \frac{e^{1/3}4m^{1/2}\mathcal{H}(F)}{|D(F)|^{1/4}(\log m)^{2\gamma}}.\end{aligned}$$

Now by Theorem 4 (setting δ there to be 2γ here), with fewer than $8 + 136(\log m)^{\varepsilon}$ possible exceptions, the primitive solutions to (1) are in certain sublattices Λ of determinant $D_F(m)$, each such sublattice can contain at most one pair $\pm\mathbf{x}$ of primitive solutions, and this only if the first minima of the sublattice satisfies $\|\mathbf{x}\| = \lambda_1 \le (\det(\Lambda))^{1/2}(\log\det(\Lambda))^{-\gamma}$.

Therefore, with no exceptions if $d \ge 5$ and at most $8 + 136(\log m)^{\varepsilon}$ exceptions if $d = 4$, we are reduced to primitive solutions that are lattice points of exceptionally small length in one of our sublattices. Such sublattices are definitely atypical.

THEOREM 5. *Let m be a positive integer and N denote the number of sublattices $\Lambda \subseteq \mathbb{Z}^2$ of determinant m. If $\gamma > 0$ and N' denotes the number of sublattices Λ of determinant m with a primitive $\mathbf{x} \in \Lambda$ satisfying $\|\mathbf{x}\| \le m^{1/2}m^{-\gamma}$, then the proportion $N'/N < 6\pi m^{-2\gamma}$. If $m > 1$ and N'' denotes the* *number of sublattices* Λ *of determinant* m *with a primitive* $\mathbf{x} \in \Lambda$ *satisfying* $\|\mathbf{x}\| \le m^{1/2} \log m^{-\gamma}$, *then the proportion* $N''/N < 6\pi(\log m)^{-2\gamma}$.

Via the discussion above, we believe that the corollary to Theorem 3 in conjunction with Theorem 5 lend credence to Stewart's (and whence Schmidt's) conjecture when the degree $d \ge 5$, and together Theorems 4 and 5 lend credence to a strong form of Schmidt's conjecture when the degree $d = 4$.

Proof of Theorem 3. Recall the hypotheses

$$\begin{aligned} \big(2D_F\big(m'\big)/\pi\big)^{1+\delta} &\ge \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}, \\ 1+\delta < 1 + \frac{d-4-\varepsilon}{2} &= \frac{d-2-\varepsilon}{2}. \end{aligned} \tag{13}$$

Now any solution $\mathbf{x} \in \mathbb{Z}^2$ to (1′) with $\|\mathbf{x}\|^{d-2-\varepsilon} > (2D_F(m')/\pi)^{1+\delta}$ is necessarily ε-exceptional, since then by Lemma 2 and (13)

$$\begin{aligned} \frac{|L_i(\mathbf{x})|}{\|\mathbf{L}_i\|} &\le \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{\|\mathbf{x}\|^{d-1} |D(F)|^{1/2}} \\ &\le \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{\|\mathbf{x}\|^{1+\varepsilon} \|\mathbf{x}\|^{d-2-\varepsilon} |D(F)|^{1/2}} \\ &< \frac{1}{\|\mathbf{x}\|^{1+\varepsilon}} \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{(2D_F(m')/\pi)^{1+\delta} |D(F)|^{1/2}} \\ &\le \frac{1}{\|\mathbf{x}\|^{1+\varepsilon}} \end{aligned}$$

for some index i. If $\Lambda \subseteq \mathbb{Z}^2$ is a sublattice with $\det(\Lambda) = D_F(m')$ and we denote the successive minima of Λ with respect to the unit disk by $\lambda_1 \le \lambda_2$, then by Minkowski's theorem

$$\lambda_2^2 \ge \lambda_1 \lambda_2 \ge \big(2^2/2!\big) \frac{\det(\Lambda)}{\pi} = 2D_F\big(m'\big)/\pi. \tag{14}$$

Now suppose $\mathbf{x} \in \Lambda$ is a primitive solution to (1′) that is not ε-exceptional. By (13) and what we just showed $\|\mathbf{x}\|^2 < 2D_F(m')/\pi$, so that by (14) we must have $\|\mathbf{x}\| < \lambda_2$. This inequality uniquely determines $\mathbf{x}$ up to scalar multiple by the definition of successive minima, and since $\mathbf{x}$ is primitive we necessarily have $\|\mathbf{x}\| = \lambda_1$. This completes the proof. □

Proof of the corollary to Theorem 3. Set $m' = m_F$ and δ so that

$$\big(2D_F(m)/\pi\big)^{1+\delta} = \frac{2^{d-1} m \mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}.$$

Note that $\delta < (d-4-\varepsilon)/2$ by hypothesis, so that we may apply Theorem 3. The result follows immediately from Theorems 1 and 3. □

Our proof of Theorem 4 requires the following.

LEMMA 9. *Let* $F(X,Y)\in\mathbb{Z}[X,Y]$ *be a homogeneous polynomial of degree* $d\geq 3$ *with non-zero discriminant and content* 1. *Suppose* m *is a positive integer and* $d-2>\varepsilon>0$. *Then the number of primitive solutions* $\mathbf{x}\in\mathbb{Z}^2$ *to* (1′) *that are not* ε*-exceptional and satisfy*

$$\|\mathbf{x}\|\geq e2^{1/(d-2)}\left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{1/(d-2)}$$

is no more than

$$2d\left(1+\frac{\log(\varepsilon/(d-2-\varepsilon))+\log\log(2^{d-1}m\mathcal{H}(F)^{d-2}/|D(F)|^{1/2})-\log(d-2)}{\log(d-1)}\right).$$

Proof. For notational convenience set

$$B=\left(\frac{2^{d-1}m\mathcal{H}(F)^{d-2}}{|D(F)|^{1/2}}\right)^{1/(d-2)}.$$

As in the proof of Theorem 3, any solution $\mathbf{x}$ to (1′) is necessarily ε-exceptional if

$$\|\mathbf{x}\|>B^{(d-2)/(d-2-\varepsilon)}=B^{\varepsilon/(d-2-\varepsilon)}B.$$

Write $F(X,Y)=\prod_{i=1}^{d}L_i(X,Y)$ as a product of linear forms. According to Lemma 2, given any solution $\mathbf{x}$ to (1′) there is an index i such that

$$\frac{|L_i(\mathbf{x})|}{\|\mathbf{L}_i\|}\leq\frac{B^{d-2}}{\|\mathbf{x}\|^{d-1}}. \tag{15}$$

Fix an index i for the moment and denote the primitive solutions to (1′) in the statement of the lemma that satisfy (15) by $\pm\mathbf{x}_j$, $j=1,2,\dots$ and arrange these so that $e2^{1/(d-2)}B\leq\|\mathbf{x}_1\|\leq\|\mathbf{x}_2\|\leq\cdots$. We claim that

$$2\|\mathbf{x}_{j+1}\|\geq\|\mathbf{x}_j\|\big(\|\mathbf{x}_j\|/B\big)^{d-2} \tag{16}$$

for all indices j.

To see our claim, consider the two linear forms M_j and M_{j+1} given by $M_j(\mathbf{w})=\mathbf{x}_j\cdot\mathbf{w}$ and $M_{j+1}(\mathbf{w})=\mathbf{x}_{j+1}\cdot\mathbf{w}$. Since $\mathbf{x}_j\neq\pm\mathbf{x}_{j+1}$ are distinct primitive points by construction, these two linear forms are linearly independent. Note that $M_j(\mathbf{L}_i)=L_i(\mathbf{x}_j)$ and $M_{j+1}(\mathbf{L}_i)=L_i(\mathbf{x}_{j+1})$. According to [10, Lemma 4], we must have one of the following inequalities:

$$\frac{|L_i(\mathbf{x}_j)|}{\|\mathbf{x}_j\|}\geq\frac{\|\mathbf{L}_i\||\det(\mathbf{x}_j^{\mathrm{tr}},\mathbf{x}_{j+1}^{\mathrm{tr}})|}{2\|\mathbf{x}_j\|\|\mathbf{x}_{j+1}\|},\qquad\frac{|L_i(\mathbf{x}_{j+1})|}{\|\mathbf{x}_{j+1}\|}\geq\frac{\|\mathbf{L}_i\||\det(\mathbf{x}_j^{\mathrm{tr}},\mathbf{x}_{j+1}^{\mathrm{tr}})|}{2\|\mathbf{x}_j\|\|\mathbf{x}_{j+1}\|}.$$

This in conjunction with (15) and the hypothesis $\|\mathbf{x}_j\|\leq\|\mathbf{x}_{j+1}\|$ shows that

$$\big|\det\big(\mathbf{x}_j^{\mathrm{tr}},\mathbf{x}_{j+1}^{\mathrm{tr}}\big)\big|\leq\frac{2\|\mathbf{x}_{j+1}\|B^{d-2}}{\|\mathbf{x}_j\|^{d-1}}. \tag{17}$$

Since $\mathbf{x}_j$ and $\mathbf{x}_{j+1}$ are linearly independent integral points, the determinant here is at least 1, whence inequality (16).

Via (16) and a simple induction argument, we see that for all $j \geq 1$ we must have $\|\mathbf{x}_j\| \geq e^{(d-1)^{j-1}} 2^{1/(d-2)} B$. In particular, since $\mathbf{x}_j$ is not ε-exceptional by hypothesis we must have $e^{(d-1)^{j-1}} \leq B^{\varepsilon/(d-2-\varepsilon)}$, so that

$$j - 1 \leq \frac{\log(\varepsilon/(d-2-\varepsilon)) + \log\log B}{\log(d-1)}.$$

Considering the d possible indices i completes the proof. □

Proof of Theorem 4. Set

$$B = \frac{2^{3/2} m^{1/2} \mathcal{H}(F)}{|D(F)|^{1/4}}$$

as in the proof of Lemma 9 (with $d = 4$ now). Assuming $\varepsilon < 1$ (so that $\varepsilon/(d-2-\varepsilon) < 1$) one readily verifies via Lemma 9 that there are less than $8(1 + \log\log B/\log 3)$ primitive solutions $\mathbf{x}$ to (1) with $\|\mathbf{x}\| \geq e2^{1/2}B$ that are not ε-exceptional. For any divisor $m' \nmid D(F)$ of m, all primitive solutions $\mathbf{x}$ to (1), in particular those solutions with

$$\frac{e^{1/3} 2^{1/2} B}{(\log m)^{\delta}} \leq \|\mathbf{x}\| < e2^{1/2} B, \tag{18}$$

lie in $c_F(m')$ sublattices of determinant $D_F(m')$ by Theorem 1. We claim that any sublattice Λ with $\det(\Lambda) \geq (\log m)^{3\delta}$ can contain at most 4 pairs $\pm\mathbf{x}$ of such primitive solutions. Indeed, our form F has 4 linear factors, so given any five pairs $\pm\mathbf{x}_j$, $1 \leq j \leq 5$, of primitive solutions to (1′) there are distinct indices $j_1 \neq j_2$ with $\|\mathbf{x}_{j_1}\| \leq \|\mathbf{x}_{j_2}\|$ satisfying

$$\left|\det\left(\mathbf{x}_{j_1}^{\mathrm{tr}}, \mathbf{x}_{j_2}^{\mathrm{tr}}\right)\right| \leq \frac{2\|\mathbf{x}_{j_2}\| B^2}{\|\mathbf{x}_{j_1}\|^3}$$

by (17). If both of these solutions satisfy (18), we have

$$\begin{aligned}
\left|\det\left(\mathbf{x}_{j_1}^{\mathrm{tr}}, \mathbf{x}_{j_2}^{\mathrm{tr}}\right)\right| &\leq \frac{2\|\mathbf{x}_{j_2}\|}{\|\mathbf{x}_{j_1}\|} \frac{B^2}{\|\mathbf{x}_{j_1}\|^2} \\
&< 2e^{2/3} (\log m)^{\delta} e^{-2/3} 2^{-1} (\log m)^{2\delta} \\
&= (\log m)^{3\delta}.
\end{aligned}$$

But if both $\mathbf{x}_{j_1}$ and $\mathbf{x}_{j_2}$ are in the sublattice Λ, then this determinant is necessarily at least as large as $\det(\Lambda)$. This proves our claim, whence the estimate for the number of primitive solutions in Theorem 4. The remainder of the theorem follows exactly as in the proof of Theorem 3. □

Proof of Theorem 5. The number of sublattices $\Lambda \subseteq \mathbb{Z}^2$ with determinant m is equal to $\sum_{n|m} n$ (see [6, Section 3], for example), thus the number N of such lattices satisfies $m < N \ll m \log\log m$ (though we will only use the lower bound here). Given a positive integer m and a primitive $(x, y) \in \mathbb{Z}^2$ with $m > (\pi/2)\|(x, y)\|^2$, we claim there is exactly one sublattice $\Lambda \subseteq \mathbb{Z}^2$

with $\det(\Lambda) = m$ containing (x, y). Indeed, for such a sublattice we must have $\|(x, y)\| < \lambda_2$ by (14). This determines (x, y) up to a scalar multiple, and since (x, y) is a primitive point we must have $\|(x, y)\| = \lambda_1$ and whence $\Lambda = \mathbb{Z}(x, y) \oplus \mathbb{Z}(x', y')$ for some $(x', y') \in \mathbb{Z}^2$ with $xy' - x'y = m = \det(\Lambda)$. Since (x, y) is a primitive point, there is an $(x', y') \in \mathbb{Z}^2$ with $xy' - x'y = m$. Moreover, from elementary number theory any other such point is of the form $(x', y') + n(x, y)$ for some integer n. Therefore all such sublattices have the same basis, so that there is only one such sublattice.

Now suppose $\gamma > 0$. If $m^{2\gamma} \leq \pi/2$ then there is nothing to prove since now $6\pi \log m^{-2\gamma} > 6\pi m^{-2\gamma} \geq 12$ (clearly larger than either proportion N'/N or N''/N). Suppose that $(\pi/2)m^{1-2\gamma} < m$. By what we have shown, the number N' of sublattices $\Lambda \subseteq \mathbb{Z}^2$ with $\det(\Lambda) = m$ that contain a primitive (x, y) with $\|(x, y)\| \leq m^{1/2}m^{-\gamma}$ is equal to the number of such primitive points. Clearly this is no greater than the total number of integral points in the disk with radius $m^{1/2}m^{-\gamma}$, which in turn is no greater than $6\pi m^{1-2\gamma}$. (The number of integral points in the disk of radius $r \geq 1$ is no more than $\pi(r + \sqrt{2})^2 \leq 6\pi r^2$.) Thus $N' \leq 6\pi m^{1-2\gamma}$ and so the proportion $N'/N < 6\pi m^{-2\gamma}$. A similar argument works to bound the proportion N''/N. $\square$

References

[1] E. Artin, *Algebraic numbers and algebraic functions*, Gordon and Breach, Basel, 1967. MR 0237460

[2] E. Bombieri and W. M. Schmidt, *On Thue's equation*, Invent. Math. **88** (1987), 69–81. MR 0877007

[3] J.-H. Evertse, *On equations in S-units and the Thue–Mahler equation*, Invent. Math. **75** (1984), 561–584.

[4] D. J. Lewis and K. Mahler, *On the representation of integers by binary forms*, Acta Arith. **6** (1961), 333–363. MR 0120195

[5] N. Saradha and D. Sharma, *The number of representations of integers by binary forms*, Publ. Math. Debrecen **85** (2014), 233–255. MR 3231517

[6] W. M. Schmidt, *Asymptotic formulæ for point lattices of bounded determinant and subspaces of bounded height*, Duke Math. J. **35** (1968), 327–339. MR 0224562

[7] W. M. Schmidt, *Diophantine approximations and diophantine equations*, Springer Lecture Notes in Mathematics, vol. 1467, Springer-Verlag, New York, 1991.

[8] C. Stewart, *On the number of solutions of polynomial congruences and Thue equations*, J. Amer. Math. Soc. **4** (1991), 793–835.

[9] A. Thue, *Über Annäherungswerte algebraischer Zahlen*, J. Reine Angew. Math. **135** (1909), 284–305.

[10] J. L. Thunder, *Decomposable form inequalities*, Ann. Math. **153** (2001), 767–804. MR 1836288

[11] J. L. Thunder, *Asymptotic estimates for the number of integer solutions to decomposable form inequalities*, Comp. Math. **141** (2005), 271–292. MR 2134267

[12] A. Weil, *Basic number theory*, Springer-Verlag, New York, 1974. MR 0427267

Jeffrey Lin Thunder, Department of Mathematics, Northern Illinois University, DeKalb, IL 60115, USA

E-mail address: jthunder@math.niu.edu