

REPRESENTATION OF INTEGERS BY SPARSE BINARY FORMS

SHABNAM AKHTARI AND PALOMA BENGOCHEA

ABSTRACT. We will give new upper bounds for the number of solutions to the inequalities of the shape $|F(x, y)| \leq h$, where $F(x, y)$ is a sparse binary form, with integer coefficients, and h is a sufficiently small integer in terms of the discriminant of the binary form F . Our bounds depend on the number of non-vanishing coefficients of $F(x, y)$. When F is “really sparse”, we establish a sharp upper bound for the number of solutions that is linear in terms of the number of non-vanishing coefficients. This work will provide affirmative answers to a number of conjectures posed by Mueller and Schmidt in [18] in special but important cases.

1. INTRODUCTION AND STATEMENTS OF THE RESULTS

Let $F(x, y)$ be a binary form of degree $n \geq 3$ with integer coefficients which is irreducible over the rationals. Let h be a positive integer. By a classical result of Thue in [22], we know that the inequality

$$(1) \quad 1 \leq |F(x, y)| \leq h$$

has at most finitely many solutions in integers x and y . Such inequalities are called *Thue’s inequalities*.

We will give upper bounds for the number of solutions to Thue’s inequalities $|F(x, y)| \leq h$, where $F(x, y)$ is a sparse polynomial, and h is sufficiently small in terms of the absolute value of the discriminant of F . Our bounds depend on the number of non-vanishing coefficients of the form F . To state our results more precisely, let us suppose that $F(x, y)$ is a form of degree $n \geq 3$ which has no more than $s + 1$ nonzero coefficients, so that

$$(2) \quad F(x, y) = \sum_{i=0}^s a_i x^{n_i} y^{n-n_i}$$

with $0 = n_0 < n_1 < \dots < n_{s-1} < n_s = n$. We refer to such forms as sparse forms or *fewnomials*.

Definition of Primitive Solutions. A pair $(x, y) \in \mathbb{Z}^2$ is called a primitive solution to inequality (1) if it satisfies the inequality and $\gcd(x, y) = 1$.

We note that by this definition $(\pm 1, 0)$ and $(0, \pm 1)$ are primitive, but, for example, if $z \neq \pm 1$, the possible solution $(z, 0)$ is not considered primitive.

2000 *Mathematics Subject Classification.* 11D45.

Key words and phrases. Sparse binary forms, Thue’s inequalities, Fewnomials, Rational approximation.

Suppose that (x_1, y_1) is a solution to the inequality (1), then there exists an integer $t \neq 0$ such that $(x_1, y_1) = (tx_2, ty_2)$ and (x_2, y_2) is a primitive solution of (1).

Throughout this manuscript, by $A \ll B$ we mean A is bounded above by B up to an explicit constant that does not depend on any of the quantities n, s, h . Similarly, we say $A = \mathcal{O}(B)$ if $A \leq \kappa B$, for an absolute constant κ . The following are our main theorems.

Theorem 1.1. *Let $F(x, y) \in \mathbb{Z}[x, y]$ be an irreducible binary form with $s+1$ nonzero coefficients, degree $n > s$, and discriminant D . Let h be an integer with*

$$(3) \quad 0 < h < \frac{|D|^{\frac{1}{8(n-1)}}}{(3n^{800 \log^2 n})^{n/2} (ns)^{2s+n}}.$$

Let $\mathcal{N}(F, n, s, h)$ be the number of primitive solutions to the inequality

$$1 \leq |F(x, y)| \leq h.$$

(i) *We have*

$$\mathcal{N}(F, n, s, h) \ll s \log s \min\left(1, \frac{1}{\log n - \log s}\right).$$

(ii) *Moreover, if $n \geq s^2$, we have*

$$\mathcal{N}(F, n, s, h) \ll s.$$

Theorem 1.2. *Let $F(x, y) \in \mathbb{Z}[x, y]$ be an irreducible binary form with $s+1$ nonzero coefficients, degree $n > s$, and discriminant D . Let h be an integer with*

$$(4) \quad 0 < h < \frac{|D|^{\frac{1}{4(n-1)}}}{10^n n^{\frac{n}{4(n-1)}}}.$$

Let $\mathcal{N}(F, n, s, h)$ be the number of primitive solutions to the inequality

$$1 \leq |F(x, y)| \leq h.$$

We have

$$\mathcal{N}(F, n, s, h) \ll \sqrt{ns}.$$

The assumptions (3) and (4) are quite strong, and are indeed helpful in our improvement of the previous bounds. Generally one cannot expect that a binary form of degree n has such a large discriminant. However, by a result of Birch and Merriman in [4], for a fixed degree n only finitely many equivalence classes of irreducible binary forms of degree n have bounded discriminant (see also [8]). So our results, while stated for a strong condition on the discriminant, namely (3) or (4), hold for almost all binary forms of a given degree.

In [17] Mueller and Schmidt obtained the upper bound

$$(5) \quad \mathcal{N}(F, n, s, h) \ll s^2 h^{2/n} (1 + \log h^{1/n}),$$

for every positive integer h . They could remove the logarithmic factor if $n \geq 4s$. One of our contributions is to remove the dependency on h for sufficiently small values of h . When h is large, one naturally expects the factor $h^{2/n}$ to appear (see [11, 24], for example). After the statement of Theorem 1 in [17], which contains the bound (5), the authors conjecture that the factor s^2 should be replaced by s (a conjecture originally due to Siegel). Our Theorem 1.1 verifies this conjecture in case $s^2 < n$ and h is small. Theorem 1.1 also improves the factor s^2 in (5) significantly, namely by a quantity smaller than $s \log s$, again for small h .

Schmidt in [18] showed, for every positive integer h , that

$$(6) \quad \mathcal{N}(F, n, s, h) \ll (ns)^{1/2} h^{2/n} (1 + \log h^{1/n}).$$

He conjectured that the logarithmic factor is unnecessary. In [23] Thunder replaced the factor $(1 + \log h^{1/n})$ in the above bound by $(1 + \log \log h / \log d)$. Using an effective result of Evertse and Györy in [8] on bounds for the height of binary forms in terms of their discriminant, Thunder reasons that Schmidt's conjecture on unnecessary of $\log h^{1/n}$ holds "essentially". Here we will remove the dependence on h in the upper bound (6), for sufficiently small h .

The problem of counting the number of solutions to Thue's inequalities $1 \leq |F(x, y)| \leq h$ for small integers h has been considered previously, for example, in [1, 9, 21], where upper bounds are of the shape $c_0 n$, where n is the degree of F and c_0 is an explicit constant. We aim for similar studies for fewnomials F . Mueller in [15] and Mueller and Schmidt in [16] established bounds for the number of solutions of $|F(x, y)| \leq h$ for binomial and trinomial forms F . These bounds are independent of h and n , provided that h is small in terms of $H(F)$, the maximum of absolute values of the coefficients of F . Based on their works on binomials and trinomials, Mueller and Schmidt conjectured in [17], provided that $h \leq H^{1 - \frac{s}{n} - \rho}$, that the number of primitive solutions of (1) is $\leq c(s, \rho)$, where $c(s, \rho)$ depends on s and ρ only. To us it feels more natural to compare the size of h with the discriminant (our methods are in sympathy with our intuition!). Our results in Theorems 1.1 and 1.2 are under the assumption that the integer h is bounded in terms of the absolute discriminant. The height and the discriminant of a binary form are connected, and our results can be stated for h bounded above by H . However, they do not imply the exact exponent of H conjectured by Mueller and Schmidt.

A very special and interesting type of fewnomials are binomials. In [20], Siegel showed that the equation $0 < |ax^n - by^n| \leq c$ has at most one primitive solution in positive integers x and y if

$$|ab|^{\frac{n}{2}-1} \geq 4 \left(n \prod_{p|n} p^{\frac{1}{p-1}} \right)^n c^{2n-2}.$$

We note that the size of c is compared to the discriminant, and not the height, of the binomial in Siegel's work. Also in [6] Evertse extended the hypergeometric method of Siegel to give striking bounds for the number of solutions to Thue equation $ax^2 - by^2 = c$. These ideas have recently been generalized in [2] for a larger family of Thue's inequalities which include binomial inequalities. In a breakthrough work [3], Bennett used a sophisticated combination of analytic methods, including the approximation tools from [20, 6], to show that the equation $ax^n - by^n = 1$, with a and b positive, has at most one solution in positive integers x, y . This is a sharp result, as the equation $(a+1)x^n - ay^n = 1$ has precisely one solution $(1, 1)$ in positive integers, for every positive integer a .

We will closely follow two fundamental works [17, 18] of Mueller-Schmidt and Schmidt. These two papers introduce different approximation methods and result in two different types of upper bounds, which are of the shapes stated in our Theorems 1.1 and 1.2. In order to discuss how we use these ideas more clearly, we organize this article in two main parts. Part I explores some ideas in [17] and [5] further and establishes the desired bounds in Theorem 1.1. Part II focuses on some extensions of results in [18] and contains the proof of Theorem 1.2. In both parts having the integer h bounded by a function of discriminant is absolutely crucial in our application of approximation methods of [17, 18]. Particularly in Part I, using the fact that the height of a binary form can also be bounded in terms of its absolute discriminant, which is invariant up to $\mathrm{GL}_2(\mathbb{Z})$ actions allows us to work with a given fewnomial and use some ideas from [5]. Simply put, we control the size of all quantities that show up in classical approximation methods by the absolute value of the discriminant.

We use two very important Lemmas established by Mueller and Schmidt, and recorded here as Propositions 4.6 and 5.1. Their paper [17] includes an interesting discussion of the Newton polygon and its applications to the distribution of the roots of polynomials with only $s+1$ roots. It turns out that the roots of such polynomial are located in not more than s fairly narrow annuli centered at the origin. This analysis is essential in establishing the extremely useful fact that every solution (x, y) of the Thue's inequality gives a good rational approximation $\frac{x}{y}$ to a root of the fewnomial that belongs to a small subset of the set of all roots. The number of elements in this subset is estimated by s , as opposed to the general case, where one has to take into account all n roots of a binary form of degree n .

In our proofs in Part I (where we ultimately prove Theorem 1.1), we will need to assume that $n > 10s$. This assumption does not alter the statement of Theorem 1.1, because technically if $n \leq 10s$, the form $F(x, y)$ is not sparse and we can use a result of the first author for general Thue's inequalities in [1], where the upper bound c_0n is established for the number of solutions of Thue's inequality $|F(x, y)| \leq h$ of degree n , and under the assumption (3). This way if $n \leq 10s$, we obtain the bound c_1s for the number of primitive solutions to our inequality, where $c_1 = 10c_0$ is an explicit constant.

This manuscript is organized as follows. After recalling some basic facts and useful theorems in Section 2, we will divide the article to two general parts. Part I includes Sections 3, 4, 5, and is devoted to the proof of Theorem 1.1. Part II includes Sections 6, 7, and is devoted to the proof of Theorem 1.2. In both parts we estimate the number of solutions to our inequalities by splitting them in three or two subsets respectively: small, medium and large solutions for Part I, and small and large solutions for Part II. The definition of small and large will differ in Parts I and II. We will give these definitions and other notation in Sections 3 and 6. The estimation of the number of large solutions is not treated in details here, as some good bounds for the number of large solutions of Thue's inequalities have been established in [18] and [17]. We will define the size of solutions in a way that we can use corresponding previous results.

2. PRELIMINARIES

2.1. Discriminant, Height, and Mahler Measure. For a binary form $G(x, y)$ that factors over \mathbb{C} as

$$\prod_{i=1}^n (\alpha_i x - \beta_i y),$$

the discriminant $D(G)$ of G is given by

$$D(G) = \prod_{i < j} (\alpha_i \beta_j - \alpha_j \beta_i)^2.$$

Therefore, if we write

$$G(x, y) = c(X - \gamma_1 y) \dots (x - \gamma_n y),$$

we have

$$D(G) = c^{2(n-1)} \prod_{i < j} (\gamma_i - \gamma_j)^2.$$

The Mahler measure $M(G)$ of the form $G(x, y) = c(X - \gamma_1 y) \dots (x - \gamma_n y)$ is defined by

$$M(G) = |c| \prod_{i=1}^n \max(1, |\beta_i|).$$

Mahler [13] showed

$$(7) \quad M(G) \geq \left(\frac{|D|}{n^n} \right)^{\frac{1}{2n-2}},$$

where D is the discriminant of G .

Let $G(x, y) = a_n x^n + a_{n-1} x^{n-1} y + \dots + a_1 x y^{n-1} + a_0 y^n$. The (naive) height of G , denoted by $H(G)$, is defined by

$$(8) \quad H(G) = \max(|a_n|, |a_{n-1}|, \dots, |a_0|).$$

We have

$$(9) \quad \binom{n}{\lfloor n/2 \rfloor}^{-1} H(G) \leq M(G) \leq (n+1)^{1/2} H(G).$$

A proof of this fact can be found in [12].

2.2. $GL_2(\mathbb{Z})$ Actions and Equivalent Forms. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and define the binary form F_A by

$$F_A(x, y) = F(ax + by, cx + dy).$$

We say that two binary forms F and G are equivalent if $G = \pm F_A$ for some $A \in GL_2(\mathbb{Z})$.

Observe that for any 2×2 matrix A with integer entries

$$(10) \quad D(F_A) = (\det A)^{n(n-1)} D(F).$$

For $A \in GL_2(\mathbb{Z})$, we have that $F_{A^{-1}}(ax + by, cx + dy) = \pm F(x, y)$ and $\gcd(ax + by, cx + dy) = 1$ if and only if $\gcd(x, y) = 1$. Therefore, the number of solutions (and the number of primitive solutions) to Thue's inequalities does not change if we replace the binary form with an equivalent form. Moreover the discriminants of two equivalent forms are equal. However, $GL_2(\mathbb{Z})$ -actions do not preserve the fact that F has no more than $s + 1$ non zero coefficients. Also $GL_2(\mathbb{Z})$ -actions do not preserve the height. So the counting problem for forms of the kind (2) does change.

To estimate the number of solutions to Thue's inequalities with fewnomials, Schmidt formulates in [18] a condition that is invariant under $GL_2(\mathbb{Z})$ actions. Following Schmidt, we define a class $C(t)$ of forms of degree n as follows.

Definition of $C(t)$. The set of forms $F(x, y)$ of degree n with integer coefficients, and irreducible over \mathbb{Q} , such that for any reals $u, v \neq 0, 0$, the form

$$(11) \quad uF_x + vF_y$$

has at most t real zeros.

Note that for $n > 0$, the irreducibility of F implies that the form (11) of degree $n - 1$ is not identically zero. Also for $F \in C(t)$, the derivative $f'(z) = F_x(z, 1)$ has fewer than t real zeros. The following is Lemma 2 of [18].

Lemma 2.1. *Suppose $F(x, y)$ is irreducible of degree n , and has $s + 1$ non-vanishing coefficients. Then $F(x, y) \in C(4s - 2)$.*

In Part II, we will consider inequalities of the shape $|F(x, y)| \leq h$, for forms $F \in C(4s - 2)$.

3. PART I: STRATEGY, OUTLINE AND DEFINITIONS

Let $\mathbf{x} = (x, y)$. We define

$$|\mathbf{x}| = \max(|x|, |y|), \quad \langle \mathbf{x} \rangle = \min(|x|, |y|).$$

In order to establish our upper bounds in Theorem 1.1, we measure the size of possible solutions (x, y) of our inequality by the size of $\langle \mathbf{x} \rangle$ and $|\mathbf{x}|$.

Definition. Relative to two quantities Y_S, Y_L , which will be defined below in (15) and (16), we call a solution $(x, y) \in \mathbb{Z}^2$

$$\begin{aligned} \text{small} & \quad \text{if} \quad 0 < \langle \mathbf{x} \rangle \leq Y_S, \\ \text{medium} & \quad \text{if} \quad |\mathbf{x}| \leq Y_L \text{ and } \langle \mathbf{x} \rangle > Y_S, \\ \text{large} & \quad \text{if} \quad |\mathbf{x}| > Y_L. \end{aligned}$$

We choose the constants below to be consistent with Mueller and Schmidt's work [17]. Let

$$(12) \quad R = n^{800 \log^2 n}.$$

From (3), we have

$$(13) \quad 0 < h \leq \frac{|D|^{\frac{1}{8(n-1)}}}{(3R)^{n/2} (ns)^{2s+n}}.$$

Put

$$C = Rh(2H\sqrt{n(n+1)})^n.$$

Pick numbers a, b with $0 < a < b < 1$ and let

$$(14) \quad \lambda = \frac{2}{(1-b)\sqrt{2/(n+a^2)}}.$$

Note that if a, b were chosen sufficiently small, then

$$n - \lambda = n - (\sqrt{2n + 2a^2}/(1-b)) > 0.$$

Definitions of Y_S and Y_L . We define

$$(15) \quad Y_S = ((e^6 s)^n R^{2s} h)^{\frac{1}{n-2s}},$$

and

$$(16) \quad Y_L = (2C)^{1/(n-\lambda)} (4(H(n+1)^{1/2} e^{n/2}))^{\lambda/((n-\lambda)a^2)}.$$

By (13), we have

$$(3R)^{n/2} h \leq \frac{|D|^{\frac{1}{8(n-1)}}}{(ns)^{2s+n}},$$

which, together with (7) and (9), implies

$$(17) \quad Y_L \ll H^2.$$

Proposition 3.1. *The number of primitive small solutions of (1) with F a fewnomial, defined in (2), and h satisfying (13) is no greater than $12s + 16$.*

Proposition 3.2. *Let $\mathcal{N}_m(F, n, s, h)$ be the number of primitive medium solutions of (1) with F a fewnomial, defined in (2), h satisfying (13), and $n > 10s$. We have*

$$\mathcal{N}_m(F, n, s, h) \ll s \log s \min\left(1, \frac{1}{\log n - \log s}\right).$$

Moreover, if $n \geq s^2$, we have

$$\mathcal{N}_m(F, n, s, h) \ll s.$$

Our assumption $n > 10s$ in Proposition 3.2 is to make our approximation methods work more smoothly. If $n \leq 10s$ then a much sharper version of Theorem 1.1 will be implied by the first author's previous work [1] on general Thue's inequalities.

Proposition 3.3. *Let $\mathcal{N}_l(F, n, s, h)$ be the number of primitive large solutions of (1) with F a fewnomial, defined in (2), and h satisfying (13). We have $\mathcal{N}_l(F, n, s, h) \ll s$.*

Our definitions of R , C , λ , Y_S and Y_L are the same as in [17]. Therefore, Proposition 3.3 follows directly from [17, Prop. 1, p.211]. We prove Propositions 3.1 and 3.2 in Sections 4 and 5, respectively.

4. SMALL SOLUTIONS (I), THE PROOF OF PROPOSITION 3.1

In this section, we estimate the number of primitive solutions $\mathbf{x} = (x, y)$ to (1) with F a fewnomial, defined in (2), for which $0 \leq \langle \mathbf{x} \rangle \leq Y_S$ with h bounded by (13). We will present a number of lemmas, some variations of which have been established by others in the past. The ideas of our proofs can be found in Chapter III of Schmidt's book [19].

We first give a bound for the number of solutions such that $0 \leq y \leq Y_S$. We estimate the number of solutions with $0 \leq x \leq Y_S$ similarly. We will regard (x, y) and $(-x, -y)$ as one solution, and can assume $y \geq 0$, if we need to.

Definition of the minimal solution. Suppose that there is at least one solution to (1) with $0 \leq y \leq Y_S$ and choose (x_0, y_0) to be a solution with minimal $y \geq 0$ (if there is no small solution then Proposition 3.1 is proven). There might be more than one solution with minimal non-negative y to the Thue's inequality. We fix one of them for our entire argument in this section, and will denote it by $\mathbf{x}_0 = (x_0, y_0)$ and call it *the minimal solution*.

Definition of $L_i(x, y)$. For the binary form

$$F(x, y) = a_n(x - \alpha_1 y) \dots (x - \alpha_n y),$$

we define $L_i(x, y) = x - \alpha_i y$, for $i = 1, \dots, n$.

Let $\mathbf{x}_1 = (x_1, y_1)$ and $\mathbf{x}_2 = (x_2, y_2)$. We define

$$(18) \quad \mathcal{D}(\mathbf{x}_1, \mathbf{x}_2) = x_1 y_2 - x_2 y_1.$$

Lemma 4.1. *Suppose $\mathbf{x} = (x, y) \in \mathbb{Z}^2$, with $\gcd(x, y) \neq 0$ and $\mathbf{x} \neq \mathbf{x}_0$, satisfies $|F(x, y)| \leq h$. We have*

$$\frac{L_i(x_0, y_0)}{L_i(x, y)} - \frac{L_j(x_0, y_0)}{L_j(x, y)} = (\beta_j - \beta_i)\mathcal{D}(\mathbf{x}, \mathbf{x}_0),$$

where β_1, \dots, β_n depend on (x, y) and are such that the form

$$J(u, w) = F(x, y)(u - \beta_1 w) \dots (u - \beta_n w)$$

is equivalent to F .

Proof. This is Lemma 5 of [7], Lemma 4 of [21] and Lemma 3 of [5]. \square

Lemma 4.2. *For each $i = 1, \dots, n$, among the primitive solutions (x, y) of (1) with $0 < y \leq Y_S$, there is at most one such that $|L_i(x, y)| < 1/(2Y_S)$.*

Proof. Suppose that (x, y) and (\tilde{x}, \tilde{y}) are two of such distinct solutions with $y \leq \tilde{y}$. Then

$$\frac{1}{\tilde{y}y} \leq \left| \frac{\tilde{x}}{\tilde{y}} - \frac{x}{y} \right| \leq \left| \frac{\tilde{x}}{\tilde{y}} - \alpha_i \right| + \left| \frac{x}{y} - \alpha_i \right| < \frac{1}{yY_S},$$

so that $\tilde{y} > Y_S$, which contradicts the assumption above. \square

Suppose, without loss of generality, that

$$(19) \quad |L_1(x_0, y_0)| = \min_{1 \leq i \leq n} (|L_i(x_0, y_0)|).$$

Note that, since $|F(x_0, y_0)| \leq h$,

$$(20) \quad |L_1(x_0, y_0)| \leq h^{1/n}.$$

By Lemma 4.2, there might exist a unique primitive solution (x^*, y^*) such that $0 < y^* \leq Y_S$ and

$$|L_1(x^*, y^*)| < \frac{1}{2Y_S}.$$

We define the set

$$(21) \quad \mathbf{A} = \{(x_0, y_0), (x^*, y^*)\}.$$

We note that $1 \leq |\mathbf{A}| \leq 2$.

By Lemma 4.2, for any solution $(x, y) \neq (x^*, y^*)$ with $0 < y \leq Y_S$, we have

$$(22) \quad |L_1(x, y)| \geq \frac{1}{2Y_S}.$$

By Lemma 4.1 and (22),

$$(23) \quad \frac{|L_i(x_0, y_0)|}{|L_i(x, y)|} \geq |\beta_1 - \beta_i| |\mathcal{D}(\mathbf{x}, \mathbf{x}_0)| - 2Y_S |L_1(x_0, y_0)|,$$

where \mathcal{D} is defined in (18). For the complex conjugate $\bar{\beta}_1$ of β_1 , we also have

$$\frac{|L_i(x_0, y_0)|}{|L_i(x, y)|} \geq |\bar{\beta}_1 - \beta_i| |\mathcal{D}(\mathbf{x}, \mathbf{x}_0)| - 2Y_S |L_1(x_0, y_0)|.$$

Hence

$$\frac{|L_i(x_0, y_0)|}{|L_i(x, y)|} \geq |\operatorname{Re}(\beta_1) - \beta_i| |\mathcal{D}(\mathbf{x}, \mathbf{x}_0)| - 2Y_S |L_1(x_0, y_0)|,$$

where $\operatorname{Re}(\beta_1)$ is the real part of β_1 . Now we choose an integer $m = m(x, y)$, with $|\operatorname{Re}(\beta_1) - m| \leq 1/2$, and we obtain

$$(24) \quad \frac{|L_i(x_0, y_0)|}{|L_i(x, y)|} \geq \left(|m - \beta_i| - \frac{1}{2} \right) |\mathcal{D}(\mathbf{x}, \mathbf{x}_0)| - 2Y_S |L_1(x_0, y_0)|,$$

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

Definition of the sets \mathfrak{X}_i . Let \mathfrak{X}_i be the set of solutions $(x, y) \notin \mathbf{A}$ with $1 \leq y \leq Y_S$ and $|L_i(x, y)| \leq \frac{1}{2y}$, where $1 \leq i \leq n$.

We note that if α_i and α_j are complex conjugates then $\mathfrak{X}_i = \mathfrak{X}_j$.

Lemma 4.3. *Suppose (x_1, y_1) and (x_2, y_2) are two distinct solutions in \mathfrak{X}_i , with $0 < y_1 \leq y_2$. Then*

$$\frac{y_2}{y_1} \geq \frac{1}{2Y_S + 7/3} \max(1, |\beta_i(x_1, y_1) - m(x_1, y_1)|).$$

Proof. We follow the proof of Lemma 4 of [5]. We have that

$$(25) \quad \begin{aligned} 1 \leq |y_2 x_1 - y_1 x_2| &\leq y_1 |L_i(x_2, y_2)| + y_2 |L_i(x_1, y_1)| \\ &\leq \frac{y_1}{2y_2} + y_2 |L_i(x_1, y_1)| \\ &\leq \frac{1}{2} + y_2 |L_i(x_1, y_1)|. \end{aligned}$$

Therefore,

$$y_2 \geq \frac{1}{2|L_i(x_1, y_1)|}.$$

Combining this with (24), we get

$$(26) \quad \frac{y_2}{y_1} \geq \frac{1}{2} \left(|m - \beta_i| - \frac{1}{2} \right) \frac{|\mathcal{D}(\mathbf{x}_1, \mathbf{x}_0)|}{y_1 |L_i(x_0, y_0)|} - \frac{Y_S |L_1(x_0, y_0)|}{y_1 |L_i(x_0, y_0)|},$$

where $\beta_i = \beta_i(x_1, y_1)$ are introduced in Lemma 4.1, $m = m(x_1, y_1)$ is an integer satisfying $|\operatorname{Re}(\beta_1) - m| \leq 1/2$, with $\beta_1 = \beta_1(x_1, y_1)$.

Now, by (19), we have that $|L_1(x_0, y_0)| \leq |L_i(x_0, y_0)|$ and

$$(27) \quad \begin{aligned} \frac{|\mathcal{D}(\mathbf{x}_1, \mathbf{x}_0)|}{y_1 |L_i(x_0, y_0)|} &= \frac{\left| \frac{x_1 - x_0}{y_1 - y_0} \right|}{\left| \frac{x_0 - \alpha_i}{y_0} \right|} \geq \frac{\left| \frac{x_1 - x_0}{y_1 - y_0} \right|}{\left| \frac{x_0 - x_1}{y_0 - y_1} \right| + \frac{1}{2y_1^2}} \\ &\geq \frac{1}{1 + \frac{1}{2y_1^2 \left| \frac{x_0 - x_1}{y_0 - y_1} \right|}} \geq \frac{2}{3}, \end{aligned}$$

where the last inequality is because $\left| \frac{x_0}{y_0} - \frac{x_1}{y_1} \right| \geq \frac{1}{y_0 y_1}$ and

$$\frac{1}{2y_1^2 \left| \frac{x_0}{y_0} - \frac{x_1}{y_1} \right|} \leq \frac{y_0}{2y_1} \leq \frac{1}{2}.$$

Therefore, by (26) and (27), we have

$$\frac{y_2}{y_1} \geq \max \left(1, \left(|m - \beta_i| - \frac{1}{2} \right) \frac{1}{3} - Y_S \right) \geq \max \left(1, \frac{2}{3} (|m - \beta_i|) \frac{1}{\frac{7}{3} + 2Y_S} \right),$$

where in the second inequality we used that $\max(1, \frac{\zeta}{2} - a) \geq \frac{1}{2a+2} \max(1, \zeta)$ with $\zeta = \frac{2}{3}|m - \beta_i|$ and $a = \frac{1}{6} + Y_S$. \square

Lemma 4.4. *Suppose $(x, y) \in \mathbb{Z}^2$, and $(x, y) \notin \mathbf{A}$, with $0 < y \leq Y_S$ and $\gcd(x, y) = 1$, satisfies $|F(x, y)| \leq h$ and $|L_i(x, y)| > \frac{1}{2y}$. Then*

$$|m(x, y) - \beta_i(x, y)| \leq \frac{7}{2} + 2h^{1/n} Y_S.$$

Proof. By (24), we have

$$|m - \beta_i| \leq \left(\frac{|L_i(x_0, y_0)|}{|L_i(x, y)|} + 2Y_S |L_1(x_0, y_0)| \right) \frac{1}{|\mathcal{D}(\mathbf{x}, \mathbf{x}_0)|} + \frac{1}{2}.$$

Since $\left| \frac{x}{y} - \frac{x_0}{y_0} \right| \geq \frac{1}{yy_0}$ and we are assuming $|L_i(x, y)| > \frac{1}{2y}$, we have

$$\frac{|L_i(x_0, y_0)|}{|L_i(x, y)| |\mathcal{D}(\mathbf{x}, \mathbf{x}_0)|} \leq \frac{\left| \alpha_i - \frac{x}{y} \right| + \left| \frac{x}{y} - \frac{x_0}{y_0} \right|}{y^2 \left| \alpha_i - \frac{x}{y} \right| \left| \frac{x}{y} - \frac{x_0}{y_0} \right|} \leq \frac{1}{y^2 \left| \frac{x}{y} - \frac{x_0}{y_0} \right|} + \frac{1}{y^2 \left| \alpha_i - \frac{x}{y} \right|} \leq 3.$$

Therefore, using also that $|L_1(x_0, y_0)| \leq h^{1/n}$, by (20), and since $|\mathcal{D}(\mathbf{x}, \mathbf{x}_0)| \geq 1$, we conclude that

$$|m - \beta_i| \leq \frac{7}{2} + 2h^{1/n} Y_S. \quad \square$$

Let $(x_1, y_1) \in \mathbb{Z}^2$ be a fixed solution to the inequality $|F(x, y)| \leq h$. The form

$$J(u, w) = F(x_1, y_1)(u - \beta_1 w) \dots (u - \beta_n w)$$

is equivalent to F by Lemma 4.1. Therefore the form

$$\hat{J}(u, w) = F(x_1, y_1)(u - (\beta_1 - m)w) \dots (u - (\beta_n - m)w),$$

which is the translation of J by $m = m(x_1, y_1) \in \mathbb{Z}$, is also equivalent to F . Hence by (7),

$$(28) \quad \prod_{i=1}^n \max(1, |\beta_i(x_1, y_1) - m(x_1, y_1)|) \geq \frac{M(\hat{J})}{|F(x_1, y_1)|} \geq \frac{|D|^{\frac{1}{2n-2}}}{hn^n}.$$

Definition of \mathfrak{X} . For each set \mathfrak{X}_i , ($i = 1, \dots, n$) that is not empty, let $(x^{(i)}, y^{(i)}) \in \mathfrak{X}_i$ be the element with the largest value of y . Let \mathfrak{X} be the set of solutions of $|F(x, y)| \leq h$ that are not in \mathbf{A} and with $1 \leq y \leq Y_S$ except the elements $(x^{(1)}, y^{(1)}), \dots, (x^{(n)}, y^{(n)})$.

In order to estimate the number of elements in \mathfrak{X} , and in view of (28), we will give an upper bound for the product

$$\prod_{i=1}^n \max(1, |\beta_i(x_1, y_1) - m(x_1, y_1)|),$$

for every $(x_1, y_1) \in \mathfrak{X}$.

Lemma 4.5. *For any fixed $i \in \{1, \dots, n\}$, we have*

$$(29) \quad \prod_{(x,y) \in \mathfrak{X}} \frac{\max(1, |\beta_i(x, y) - m(x, y)|)}{Y_S(2Y_S + 7/2)} \leq 1.$$

Proof. Suppose that the set \mathfrak{X}_i is non-empty. We index the elements of \mathfrak{X}_i as

$$(x_1^{(i)}, y_1^{(i)}), \dots, (x_v^{(i)}, y_v^{(i)}),$$

so that $y_1^{(i)} \leq \dots \leq y_v^{(i)}$ (note that $(x_v^{(i)}, y_v^{(i)}) = (x^{(i)}, y^{(i)})$). By Lemma 4.3,

$$(30) \quad \frac{1}{2Y_S + 7/2} \max\left(1, \left|\beta_i(x_k^{(i)}, y_k^{(i)}) - m(x_k^{(i)}, y_k^{(i)})\right|\right) \leq \frac{y_{k+1}^{(i)}}{y_k^{(i)}}$$

for $k = 1 \dots, v - 1$. Therefore,

$$(31) \quad \prod_{(x,y) \in \mathfrak{X} \cap \mathfrak{X}_i} \frac{1}{2Y_S + 7/2} \max(1, |\beta_i(x, y) - m(x, y)|) \leq Y_S.$$

For any solution $(x, y) \in \mathfrak{X}$, with $1 \leq y \leq Y_S$, that does not belong to \mathfrak{X}_i , by Lemma 4.4, we have

$$(32) \quad \frac{\max(1, |\beta_i(x, y) - m(x, y)|)}{2h^{1/n}Y_S + 7/2} \leq 1.$$

This, together with (31), completes the proof of Lemma. \square

Next we will use the following striking result from [17] to establish inequalities similar to (29) for the solutions $(x^{(i)}, y^{(i)})$ which, by definition, do not belong to \mathfrak{X} .

Proposition 4.6. *Let F be a fewnomial, defined in (2). There is a set \mathcal{S} of roots α_i of $F(x, 1)$ with $|\mathcal{S}| \leq 6s + 4$ such that for any real ζ ,*

$$\min_{\alpha_\ell \in \mathcal{S}} |\zeta - \alpha_\ell| \leq R \min_{1 \leq i \leq n} |\zeta - \alpha_i|.$$

Proof. This is Lemma 7 of [17]. \square

Let

$$\mathbf{S}_1 = \mathbf{S} \cup \{\alpha_1\},$$

where \mathbf{S} is the set in Proposition 4.6, and α_1 is the fixed root associated to the minimal solution, for which the inequality (19) is satisfied. Proposition 4.6 implies that

$$|\mathbf{S}_1| \leq 6s + 5.$$

Let $\mathbf{S}_1 = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$, with $1 \leq t \leq 6s + 5$.

Recall that we denote by $(x^{(i)}, y^{(i)})$ the element in \mathfrak{X}_i with the largest value of y . Suppose $(x^{(i)}, y^{(i)}) \in \mathfrak{X}_i \setminus \{\mathfrak{X}_1 \cup \dots \cup \mathfrak{X}_{i-1}\}$, for $t < i \leq n$.

By Proposition 4.6, there exists $\ell \in \{1, \dots, t\}$ such that

$$|L_i(x^{(i)}, y^{(i)})| \geq \frac{|L_\ell(x^{(i)}, y^{(i)})|}{R} \geq \frac{1}{2y^{(i)}R},$$

where the last inequality is because $(x^{(i)}, y^{(i)}) \notin \mathfrak{X}_\ell$. Combining this with (24), we obtain

$$\left| m(x^{(i)}, y^{(i)}) - \beta_i(x^{(i)}, y^{(i)}) \right| \leq \frac{2|L_i(x_0, y_0)|y^{(i)}R + 2Y_S|L_1(x_0, y_0)|}{|\mathcal{D}(\mathbf{x}^{(i)}, \mathbf{x}_0)|} + \frac{1}{2}.$$

Using (27) and $|L_1(x_0, y_0)| \leq h^{1/n}$, we obtain

$$\left| m(x^{(i)}, y^{(i)}) - \beta_i(x^{(i)}, y^{(i)}) \right| \leq 3R + 2h^{1/n}Y_S + \frac{1}{2}.$$

This, together with (13) and (15), implies that

$$\begin{aligned} (33) \quad & \frac{|m(x^{(i)}, y^{(i)}) - \beta_i(x^{(i)}, y^{(i)})|}{Y_S(2Y_S + 7/2)} \\ & \leq \frac{3R + 2h^{1/n}Y_S + \frac{1}{2}}{Y_S(2Y_S + 7/2)} \\ & < |D|^{\frac{1}{4n(n-1)}}. \end{aligned}$$

Definition of the set \mathfrak{X}^+ . Let $\mathfrak{X}^+ = \mathfrak{X} \cup \{(x^{(i)}, y^{(i)})\}_{t < i \leq n}$.

By (28), (29), (33) and Lemma 4.4, we have that

$$\begin{aligned} (34) \quad & \left(\frac{|D|^{\frac{1}{2n-2}}}{h[nY_S(2Y_S + 7/2)]^n} \right)^{|\mathfrak{X}^+|} \\ & \leq \prod_{(x,y) \in \mathfrak{X}^+} \frac{1}{[Y_S(2Y_S + 7/2)]^n} \prod_{i=1}^n \max(1, |\beta_i(x_1, y_1) - m(x_1, y_1)|) \\ & < \left(|D|^{\frac{1}{4n(n-1)}} \right)^{n-t}. \end{aligned}$$

By (13) and (15), we have

$$\begin{aligned}
h[nY_S(2Y_S + 7/2)]^n &< h(2n)^n (Y_S^2 + 2Y_S)^n \\
&< (2n)^n h(Y_S + 1)^{2n} = (2n)^n h\left(\left((e^6 s)^n R^{2s} h\right)^{\frac{1}{n-2s}} + 1\right)^{2n} \\
&< (2n)^n h^{1+\frac{2n}{n-2s}} \left(\left((e^6 s)^n R^{2s}\right)^{\frac{1}{n-2s}} + 1\right)^{2n} \\
&< (2n)^n \left(\frac{|D|^{\frac{1}{8(n-1)}}}{(3R)^{n/2}(ns)^{2s+n}}\right)^{1+\frac{2n}{n-2s}} \left(\left((e^6 s)^n R^{2s}\right)^{\frac{1}{n-2s}} + 1\right)^{2n}.
\end{aligned}$$

Since we assumed $10s < n$, we have the following inequalities for the exponents of $|D|$, R , s in the last expression above.

$$\begin{aligned}
\frac{1}{8n(n-1)} \left(1 + \frac{2n}{n-2s}\right) &< \frac{1}{8n(n-1)} \left(\frac{7}{2}\right), \\
\frac{4ns}{n-2s} &< \frac{n}{2} \left(1 + \frac{2n}{n-2s}\right)
\end{aligned}$$

and

$$\frac{2n^2}{n-2s} < (2s+n) \left(1 + \frac{2n}{n-2s}\right).$$

Therefore,

$$(35) \quad h[nY_S(2Y_S + 7/2)]^n < |D|^{\frac{7}{16(n-1)}}.$$

By (34) and (35), we have

$$\left(\frac{|D|^{\frac{1}{2n-2}}}{|D|^{\frac{7}{16(n-1)}}}\right)^{|\mathfrak{X}^+|} < \left(|D|^{\frac{1}{8n(n-1)}}\right)^n.$$

Therefore,

$$|\mathfrak{X}^+| < \frac{\frac{1}{4(n-1)} \log |D|}{\frac{1}{16(n-1)} \log |D|} = 4.$$

The solutions (x, y) with $0 < y \leq Y_S$ are either in $\mathfrak{X}^+ \cup \mathbf{A}$ (see (21) for definition) or they are among possible $(x^{(1)}, y^{(1)}), \dots, (x^{(t)}, y^{(t)})$. Counting the solutions in the set \mathbf{A} , and by Proposition 4.6, we see that the total number of solutions with $0 \leq y \leq Y_S$ is no greater than $6s + 6 + 3 + 3$. The number of solutions of (1) such that $0 \leq x \leq Y_S$ can be estimated in a similar way, by considering the form

$$F(x, y) = a_0(y - \gamma_1 x) \cdots (y - \gamma_n y)$$

and putting $L_i(x, y) = y - \gamma_i x$. Here $\gamma_1, \dots, \gamma_n$ are the roots of the polynomial $F(1, y)$. We conclude that the number of solutions with $0 \leq x \leq Y_S$ is no greater than $6s + 12$.

5. MEDIUM SOLUTIONS (I), PROOF OF PROPOSITION 3.2

We divide the interval $[Y_S, Y_L]$ into $N + 1$ subintervals, where Y_S and Y_L are defined in (15) and (16) and N depends on s and is defined below. We will show there are only few solutions (x, y) with y in each of these subintervals. In this section we will assume $n > 10s$. We define a positive integer $N = N(n, s)$ as follows.

If $n \geq s^2$, we put $N = 2$.

If $n < s^2$, we choose $N \in \mathbb{N}$ such that

$$(36) \quad 10s^{1+\frac{1}{N}} \leq n \leq 10s^{1+\frac{1}{N-1}}.$$

If $n < s^2$, we have

$$(37) \quad N \leq \frac{\log s}{\log n - \log s}.$$

For $\ell = 1, \dots, N$, we define

$$Y_\ell = Y_S H^{\frac{1}{s^{1-(\ell-1)/N}}},$$

where H the height of F , defined in (8). We put

$$Y_0 = Y_S \quad \text{and} \quad Y_{N+1} = Y_L.$$

We will use the following important result achieved in [17].

Proposition 5.1. *There is a set \mathbf{T} of roots of $F(x, 1)$ and a set \mathbf{T}^* of roots of $F(1, y)$, both with cardinalities at most $6s + 4$, such that any solution (x, y) of (1) with $\langle \mathbf{x} \rangle \geq Y_S$ either has*

$$(38) \quad \left| \alpha - \frac{x}{y} \right| < \frac{R(ns)^2}{H^{(1/s)-(1/n)}} \left(\frac{(4e^3 s)^n h}{y^n} \right)^{1/s}$$

with some $\alpha \in \mathbf{T}$, or has

$$(39) \quad \left| \alpha^* - \frac{y}{x} \right| < \frac{R(ns)^2}{H^{(1/s)-(1/n)}} \left(\frac{(4e^3 s)^n h}{x^n} \right)^{1/s}$$

for some $\alpha^* \in \mathbf{T}^*$.

Proof. This is Lemma 17 of [17]. □

Let $\alpha \in \mathbf{T}$. For $\ell \in \{0, \dots, N\}$, let $(x_1, y_1), \dots, (x_{w_\ell}, y_{w_\ell})$ be the primitive solutions of our inequality, with $Y_\ell < y_i \leq Y_{\ell+1}$, satisfying (38) and ordered so that

$$Y_\ell < y_1 \leq \dots \leq y_{w_\ell} \leq Y_{\ell+1}.$$

By (38), we have that

$$\frac{1}{y_i y_{i+1}} \leq \left| \frac{x_{i+1}}{y_{i+1}} - \frac{x_i}{y_i} \right| \leq \frac{K}{H^{(1/s)-(1/n)} y_i^{\frac{n}{s}}},$$

with

$$K = 2R(ns)^2 (4e^3 s)^{n/s} h^{1/s}.$$

Therefore, for solutions (x, y) with $y \in [Y_\ell, Y_{\ell+1}]$, we have

$$(40) \quad y_{i+1} \geq K^{-1} H^{\frac{1}{s}-\frac{1}{n}} y_i^{\frac{n}{s}-1} \geq K^{-1} H^{\frac{1}{s}-\frac{1}{n}} Y_\ell^{\frac{n}{s}-2} y_i.$$

First we will give an estimate for the number of primitive solutions in the first subinterval $[Y_0, Y_1]$. We have $Y_0 = Y_S$. By the definition (15) of Y_S , we have

$$(41) \quad K^{-1} Y_S^{\frac{n}{s}-2} \geq 1.$$

For $\ell = 0$, we have by (40) and (41) that $y_{i+1} \geq H^{\frac{1}{s}-\frac{1}{n}} y_i$, so $y_{w_0} \geq (H^{\frac{1}{s}-\frac{1}{n}})^{(w_0-1)} y_1$. Therefore, we have

$$Y_1 \geq y_{w_0} \geq (H^{\frac{1}{s}-\frac{1}{n}})^{(w_0-1)} Y_0,$$

and

$$w_0 - 1 \leq \frac{\log \frac{Y_1}{Y_0}}{(\frac{1}{s} - \frac{1}{n}) \log H} \leq \frac{10}{9},$$

since $\log \frac{Y_1}{Y_0} = \frac{1}{s} \log H$ and $n > 10s$.

For $1 \leq \ell < N$, by (40) and (41) we have that

$$y_{i+1} \geq K^{-1} H^{\frac{1}{s}-\frac{1}{n}} Y_S^{\frac{n}{s}-2} H^{\frac{n/s-2}{s^{1-(\ell-1)/N}}} y_i \geq H^{\frac{n}{s^2-(\ell-1)/N} - \frac{2}{s^{1-(\ell-1)/N}}} y_i.$$

Therefore,

$$y_{w_\ell} \geq H^{\left(\frac{n}{s^2-(\ell-1)/N} - \frac{2}{s^{1-(\ell-1)/N}}\right)(w_\ell-1)} y_1,$$

and since $Y_\ell < y_1 < y_{w_\ell} \leq Y_{\ell+1}$, we have

$$w_\ell - 1 \leq \frac{\log \frac{Y_{\ell+1}}{Y_\ell}}{\left(\frac{n}{s^2-(\ell-1)/N} - \frac{2}{s^{1-(\ell-1)/N}}\right) \log H}.$$

For $\ell < N$, since $\log \frac{Y_{\ell+1}}{Y_\ell} < \frac{1}{s^{1-\ell/N}} \log H$ and $n \geq 10s^{1+1/N}$,

$$w_\ell - 1 \leq \frac{1}{\frac{n}{s^{1+1/N}} - \frac{2}{s^{1/N}}} < 1.$$

For $\ell = N$, we have $\log Y_{\ell+1} = \log Y_L \ll 2 \log H$, by (17), and therefore we get

$$w_N - 1 \ll \frac{2}{\frac{n}{s^{1+1/N}} - \frac{2}{s^{1/N}}} < 2.$$

We conclude that the number of primitive medium solutions of (38) for each $\alpha \in \mathbf{T}$ is $\mathcal{O}(N)$. In a similar way, the number of primitive medium solutions of (39) for each $\alpha^* \in \mathbf{T}^*$ is $\mathcal{O}(N)$. By Proposition 5.1, the number of primitive medium solutions of (1) is $\mathcal{O}(N)$. We obtain Proposition 3.2, as by our definition we have (37) for $n < s^2$, and $N = 2$ when $n \geq s^2$.

6. PART II: STRATEGY, OUTLINE AND DEFINITIONS

We will consider binary forms $F(x, y) \in C(4s-2)$ (see (11), for definition). By Lemma 2.1, our discussion for such forms implies Theorem 1.2.

Definition of Normalized and Reduced Forms. Suppose $0 \leq \mathfrak{a} \leq \mathfrak{b}$. The number of primitive solutions to the inequality

$$\mathfrak{a} \leq |F(x, y)| \leq \mathfrak{b}$$

remains unchanged if we replace F by one of its $\mathrm{GL}_2(\mathbb{Z})$ -equivalent forms. Moreover, if the inequality $\mathfrak{a} \leq |F(x, y)| \leq \mathfrak{b}$ has at least one primitive solution (x_1, y_1) , there is an $A \in \mathrm{GL}_2(\mathbb{Z})$ with $A^{-1}(x_1, y_1)^{\mathrm{tr}} = (1, 0)$, so that

$$\mathfrak{a} \leq |F_A(1, 0)| \leq \mathfrak{b}.$$

So in order to estimate the number of primitive solutions to the above inequality, we may restrict our attention to *normalized forms* for which the leading coefficient a_0 has

$$\mathfrak{a} \leq |a_0| \leq \mathfrak{b}.$$

We will say that a form F is *reduced* if it is normalized and has the smallest Mahler measure among all normalized forms equivalent to F .

Schmidt worked with reduced forms to establish the results in [18], in particular, his Lemma 4 is of special importance here (see our Lemma 7.1 and its implications).

Every primitive solution of the inequality $1 \leq |F(x, y)| \leq h$ is either a solution of

$$(42) \quad h^{1/2} < |F(x, y)| \leq h$$

or

$$(43) \quad 1 \leq |F(x, y)| \leq h^{1/2}.$$

To obtain our desired bounds, we will need to assume that the form F in (42) is normalized with respect to $(\mathfrak{a}, \mathfrak{b}) = (h^{1/2}, h)$ and the form F in (43) is normalized with respect to $(\mathfrak{a}, \mathfrak{b}) = (1, h^{1/2})$. We will have two equivalent, but not identical forms in each of the inequalities. We will show that the number of primitive solutions to each of these two inequalities is $\mathcal{O}(ns)$, provided that h satisfies

$$(44) \quad h < \frac{|D|^{\frac{1}{4(n-1)}}}{10^n n^{\frac{n}{4(n-1)}}}.$$

From now on, we consider the inequality

$$(45) \quad \mathfrak{a} \leq |F(x, y)| \leq \mathfrak{b}$$

and assume that F is reduced with respect to $(\mathfrak{a}, \mathfrak{b})$,

$$(46) \quad \mathfrak{b} \leq h,$$

and

$$(47) \quad \frac{\mathfrak{b}}{\mathfrak{a}} \leq h^{1/2}.$$

These assumptions are necessary in our estimation of linear forms in Section 7. Namely, the assumption (46) is important for our estimates in the proof of Lemma 7.2, and the assumption (45) is essential for Lemma 7.1 to hold.

We denote by M the Mahler measure $M(F)$ of the reduced form F .

By (7), our assumption (44) implies

$$(48) \quad h < \frac{M^{1/2}}{10^n},$$

and

$$(49) \quad M > 100^n h^2.$$

Similar to Schmidt's work in [18], we define

$$(50) \quad Q := \frac{M}{h} \geq 100^n h.$$

If (x, y) is a solution to our inequality, we will assume without loss of generality that $y > 0$ (recall that we count (x', y') and $(-x', -y')$ as one solution and we have at most one primitive solution with $y = 0$). We take

$$Y'_S = M^2.$$

Definition of small and large solutions. We call a solution $(x, y) \in \mathbb{Z}^2$ small if $0 < y \leq Y'_S$. We call a solution $(x, y) \in \mathbb{Z}^2$ large if $Y'_S < y$.

We will prove the following in the remaining of the manuscript.

Proposition 6.1. *Let $\mathcal{N}_1(F, n, s, h)$ be the number of small solutions of $1 \leq |F(x, y)| \leq h$, where $F(x, y) \in C(4s - 2)$ of degree n is reduced. Assume that h satisfies (44), with (45) and (46). We have $\mathcal{N}_1(F, n, s, h) \ll \sqrt{ns}$.*

For $n \gg 1$, h satisfying (44), and our quantity $Y'_S = M^2$, we have

$$Y'_S > (2h(2n^{\frac{1}{2}}M)^n)^{\frac{1}{n-\lambda}} (4(Me^{\frac{n}{2}})^{\frac{1}{a^2}})^{\frac{\lambda}{n-\lambda}},$$

where the right hand side is the quantity defined by Schmidt in [18, eq. 4.6] to distinguish between small and large solutions, where a and λ are defined as in (14). Therefore we may apply Schmidt's upper bound for the number of large solutions to our inequalities $\mathbf{a} \leq |F(x, y)| \leq \mathbf{b}$. We note here that in [18], no restriction on h is assumed, however our assumption (44) results in having the above inequality for Y'_S hold.

Proposition 6.2. *Let $\mathcal{N}_2(F, n, s, h)$ be the number of large solutions of $1 \leq |F(x, y)| \leq h$, where $F(x, y) \in C(4s - 2)$ has degree n . Assume that h satisfies (44). We have $\mathcal{N}_2(F, n, s, h) \ll \sqrt{ns}$.*

Proof. This is proven in [18]. See the discussion in the beginning of page 247 of [18], which results in Theorems 3 and 4 of [18]. \square

We conclude this section by recalling the trivial but important fact that in inequalities (42) and (43) the discriminant D is fixed, but the Mahler measure (and therefore the definition of small and large solutions) varies.

7. SMALL SOLUTIONS (II), PROOF OF PROPOSITION 6.1

We first give an upper bound for the number of solutions such that $0 < y \leq Y'_S$.

7.1. Estimation of linear forms. Let $0 \leq \mathfrak{a} \leq \mathfrak{b}$, with (45) and (46). Suppose that $F(x, y)$ belongs to the class $C(4s - 2)$, is reduced and satisfies (50).

We have

$$F(x, y) = a_0(x - \alpha_1 y) \dots (x - \alpha_n y),$$

where $\alpha_1, \dots, \alpha_n$ are the roots of the polynomial $F(x, 1)$, and put

$$L_i(x, y) = x - \alpha_i y$$

for $i = 1, \dots, n$. The following is Lemma 4 of [18]. We present its short proof here to clarify the definition of Q in (50) and more importantly the importance of the assumption $\mathfrak{b} \leq h$.

Lemma 7.1. *Suppose $G(x, y) = b_0(x - \beta_1 y) \dots (x - \beta_n y)$ is normalized and equivalent to the reduced form F , with (45), (46) and (47), and let*

$$(51) \quad \eta_i = |\beta_i - m| + 1 \quad (i = 1, \dots, n)$$

where m is an integer. Then

$$(52) \quad \eta_1 \dots \eta_n > Q,$$

where Q is given in (50).

Proof. The form

$$\hat{G}(x, y) = G(x + my, y) = b_0 \prod_{i=1}^n (x + (m - \beta_i)y)$$

is also normalized and equivalent to both G and F . Since F is reduced

$$M(F) \leq M(\hat{G}) < |b_0| \eta_1 \dots \eta_n.$$

Our proof is complete, since G is reduced and $|b_0| \leq h$. □

Our next Lemma is a modified version of Lemma 5 of [18].

Lemma 7.2. *Suppose (x_0, y_0) and (x, y) are linearly independent primitive integer points that satisfy $\mathfrak{a} \leq |F(x, y)| \leq \mathfrak{b}$, with (45), (46) and (47). Then there are numbers ψ_1, \dots, ψ_n satisfying*

$$(53) \quad \psi_i = 0 \quad \text{or} \quad \frac{1}{2n} \leq \psi_i \leq 1,$$

with

$$(54) \quad \sum_{i=1}^n \psi_i \geq \frac{1}{2},$$

such that

$$(55) \quad \left| \frac{L_i(x_0, y_0)}{L_i(x, y)} \right| \geq \left(Q^{\psi_i} - \frac{3}{2} - h^{1/2n} \right) |x_0 y - x y_0|,$$

for $i \in \{1, \dots, n\}$.

Proof. Pick $(x', y') \in \mathbb{Z}^2$ with $x'y - xy' = 1$, so that (x', y') and (x, y) form a basis for \mathbb{Z}^2 . We may write $(x_0, y_0) = a(x, y) + b(x', y')$. Then

$$x_0 y - x y_0 = b(x'y - xy') = b.$$

Therefore,

$$(56) \quad \frac{L_i(x_0, y_0)}{L_i(x, y)} = a + (x_0 y - x y_0) \frac{L_i(x', y')}{L_i(x, y)} = a - (x_0 y - x y_0) \beta_i,$$

for $i \in \{1, \dots, n\}$. (We define β_i by the second equation above.) Set

$$G(v, w) := F(v(x, y) + w(x', y')),$$

so that G is equivalent to F , and G is normalized (recall that (x, y) and (x', y') are fixed and $x'y - xy' = 1$). We have

$$\begin{aligned} G(v, w) &= a_0 \prod_{i=1}^n L_i(v(x, y) + w(x', y')) \\ &= a_0 \prod_{i=1}^n (v L_i(x, y) + w L_i(x', y')) \\ &= b_0 \prod_{i=1}^n \left(v + \frac{L_i(x', y')}{L_i(x, y)} w \right) \\ &= b_0 \prod_{i=1}^n (v - \beta_i w) \end{aligned}$$

with $b_0 = F(x, y)$. Note that

$$\left| \frac{F(x_0, y_0)}{F(x, y)} \right| \leq h^{1/2}.$$

We may assume that

$$\left| \frac{L_n(x_0, y_0)}{L_n(x, y)} \right| = \min_i \left| \frac{L_i(x_0, y_0)}{L_i(x, y)} \right|,$$

so that

$$\left| \frac{L_n(x_0, y_0)}{L_n(x, y)} \right| \leq h^{\frac{1}{2n}}.$$

By (56), we have

$$(57) \quad |a - (x_0 y - x y_0) \beta_n| \leq h^{\frac{1}{2n}},$$

and

$$|a - (x_0 y - x y_0) \beta| \leq h^{\frac{1}{2n}},$$

where β is the real part of β_n . Now let m be an integer with

$$|m - \beta| \leq \frac{1}{2}$$

and define η_1, \dots, η_n by (51), so that (52) holds by Lemma 7.1. We define

$$\eta'_i = \begin{cases} Q & \text{if } \eta_i \geq Q, \\ \eta_i & \text{if } Q^{1/2n} \leq \eta_i < Q, \\ 1 & \text{if } \eta_i < Q^{1/2n}. \end{cases}$$

We note that since $\eta_1 \dots \eta_n > Q$, we have, by the definition, that

$$(58) \quad \eta'_1 \dots \eta'_n \geq Q^{1/2}.$$

Now we define the numbers ψ_i , for $i \in \{1, \dots, n\}$, as follows

$$\eta'_i = Q^{\psi_i}.$$

Clearly ψ_i satisfy (53) and (54). By (56) and (57),

$$\begin{aligned} \left| \frac{L_i(x_0, y_0)}{L_i(x, y)} \right| &= |(\beta - \beta_i)(x_0 y - x y_0) + a - (x_0 y - x y_0)\beta| \\ &\geq |\beta - \beta_i| |x_0 y - x y_0| - h^{\frac{1}{2n}} \geq \left(|m - \beta_i| - \frac{1}{2} \right) |x_0 y - x y_0| - h^{\frac{1}{2n}} \\ &\geq \left(|m - \beta_i| - \frac{1}{2} - h^{\frac{1}{2n}} \right) |x_0 y - x y_0| \\ &= \left(\eta_i - \frac{3}{2} - h^{\frac{1}{2n}} \right) |x_0 y - x y_0|. \end{aligned}$$

Since $\eta_i \geq \eta'_i = Q^{\psi_i}$, the proof of the lemma is completed. \square

The following is a modified version of Lemma 6 in [18]:

Lemma 7.3. *Suppose (x, y) is primitive, with $y > 0$, $\mathbf{a} \leq |F(x, y)| \leq \mathbf{b}$, with (45), (46) and (47). Then there are numbers $\psi_i = \psi_i(x, y)$ ($i = 1, \dots, n$), which satisfy (53) and (54), such that*

$$(59) \quad |L_i(x, y)| < \frac{1}{Q^{\psi_i/2} y}$$

for each i with $\psi_i > 0$.

Proof. We first note, by definition, that for i with $\psi_i > 0$, we have $\psi_i \geq \frac{1}{2n}$, and by (50),

$$Q^{\psi_i} \geq Q^{\frac{1}{2n}} > 3 + 2h^{1/2n}.$$

Therefore,

$$Q^{\psi_i} - \frac{3}{2} - h^{1/2n} \geq \frac{1}{2} Q^{\psi_i} \geq Q^{\psi_i/2}$$

since, by (50) again, $Q \geq 4^{2n}$. We now apply Lemma 7.2 with $(x_0, y_0) = (1, 0)$. \square

7.2. Counting Small Solutions. We define Φ_i ($i = 1, \dots, n$) by

$$(60) \quad \begin{cases} \Phi_i = 0 & \text{when } |\operatorname{Im} \alpha_i| > 1, \\ M^{-\Phi_i} = |\operatorname{Im} \alpha_i| & \text{when } 0 < |\operatorname{Im} \alpha_i| \leq 1, \\ \Phi_i = +\infty & \text{when } \alpha_i \text{ is real.} \end{cases}$$

Lemma 7.4. *Let \mathfrak{Y} be the set of primitive integer points satisfying*

$$\mathfrak{a} \leq |F(x, y)| \leq \mathfrak{b},$$

with $\frac{\mathfrak{b}}{\mathfrak{a}} \leq h^{\frac{1}{2}}$, and $0 < y \leq Y'_S$. Then for $i = 1, \dots, n$,

$$(61) \quad \sum_{x \in \mathfrak{Y}} \psi_i(x, y) < 10 \min(1, \Phi_i) \frac{\log M}{\log Q}.$$

Proof. We follow the proof of Lemma 7 in [18]. For a fixed i , let $(x_1, y_1), \dots, (x_\nu, y_\nu)$ be the elements of \mathfrak{Y} with $\psi_i > 0$, ordered such that $y_1 \leq \dots \leq y_\nu$. By (59), we have

$$|\operatorname{Im} \alpha_i| \leq \left| \alpha_i - \frac{x_j}{y_j} \right| < \frac{1}{Q^{\psi_i(x_j, y_j)/2} y_j^2}$$

for $j = 1, \dots, \nu$. First we conclude that $|\operatorname{Im} \alpha_i| \leq 1$, so that

$$(62) \quad |\operatorname{Im} \alpha_i| = M^{-\Phi_i}$$

(with $M^{-\infty} = 0$). So we have

$$M^{-\Phi_i} < \frac{1}{Q^{\psi_i(x_j, y_j)/2} y_j^2}.$$

Therefore, for every $(x_j, y_j) \in \mathfrak{Y}$, we have

$$(63) \quad y_j < M^{\Phi_i/2}, \quad \psi_i(x_j, y_j) < 2 \frac{\log M}{\log Q} \Phi_i.$$

In particular,

$$y_\nu \leq Y_i,$$

where

$$(64) \quad Y_i := \min(Y'_S, M^{\Phi_i/2}).$$

Now let us suppose that $\nu > 1$ and $1 \leq j < \nu$. We have

$$\left| \alpha_i - \frac{x_j}{y_j} \right| < \frac{1}{Q^{\psi_i(x_j, y_j)/2} y_j^2},$$

and

$$\left| \alpha_i - \frac{x_{j+1}}{y_{j+1}} \right| < \frac{1}{Q^{\psi_i(x_{j+1}, y_{j+1})/2} y_{j+1}^2}.$$

So we have

$$\begin{aligned}
 1 &\leq |x_j y_{j+1} - x_{j+1} y_j| \\
 &= \left| (y_j y_{j+1}) \left(\frac{x_j}{y_j} - \alpha_i + \alpha_i - \frac{x_{j+1}}{y_{j+1}} \right) \right| \\
 &< \frac{y_{j+1}}{y_j Q^{\psi_i(x_j, y_j)/2}} + \frac{y_j}{y_{j+1} Q^{\psi_i(x_{j+1}, y_{j+1})/2}} \\
 &< \frac{y_{j+1}}{y_j Q^{\psi_i(x_j, y_j)/2}} + \frac{1}{3},
 \end{aligned}$$

since $\frac{y_j}{y_{j+1}} \leq 1$ and $Q^{\psi_i(x_{j+1}, y_{j+1})/2} \geq Q^{\frac{1}{4n}} > 3$ by (50). Therefore, we obtain the following *gap principle*.

$$y_{j+1} > \frac{2}{3} Q^{\psi_i(x_j, y_j)/2} y_j > Q^{\psi_i(x_j, y_j)/4} y_j \quad (1 \leq j < \nu),$$

by using (50) once again.

Applying the above gap principle repeatedly, and by the definition of Y_i in (64), we have

$$Q^{\frac{1}{4}(\psi_i(x_1, y_1) + \dots + \psi_i(x_{\nu-1}, y_{\nu-1}))} < y_\nu \leq Y_i$$

and consequently,

$$\sum_{j=1}^{\nu-1} \psi_i(x_j, y_j) < 4 \frac{\log Y_i}{\log Q}.$$

For our choice of $Y'_S = M^2$,

$$\log Y'_S = 2 \log M,$$

and therefore

$$\sum_{j=1}^{\nu-1} \psi_i(x_j, y_j) < 4 \min \left(2, \frac{\Phi_i}{2} \right) \frac{\log M}{\log Q} \leq 8 \min (1, \Phi_i) \frac{\log M}{\log Q}.$$

Now to estimate $\sum_{j=1}^{\nu} \psi_i(x_j, y_j)$, we only need to estimate $\psi_i(x_\nu, y_\nu)$. By (53) and (63), we have

$$\psi_i(x_\nu, y_\nu) \leq \min \left(1, 2\Phi_i \frac{\log M}{\log Q} \right) \leq 2 \min (1, \Phi_i) \frac{\log M}{\log Q}.$$

So we conclude the assertion of the Lemma. \square

Now we note that the number of small solutions to $\mathbf{a} \leq |F(x, y)| \leq \mathbf{b}$ is equal to $\sum_{x \in \mathfrak{Y}} 1$, and by (54) and (50), we have

$$\begin{aligned} \sum_{x \in \mathfrak{Y}} 1 &\leq 2 \sum_{i=1}^n \sum_{x \in \mathfrak{Y}} \psi_i \\ &\leq 20 \sum_{i=1}^n \min(1, \Phi_i) \frac{\log M}{\log Q} \\ &\leq 40 \sum_{i=1}^n \min(1, \Phi_i). \end{aligned}$$

In the next subsection we will show that

$$(65) \quad \sum_{i=1}^n \min(1, \Phi_i) \ll \sqrt{sn}$$

This will complete the proof of Proposition 6.1.

7.3. The clustering of roots with small imaginary parts. In order to utilize a powerful result of Schmidt in [18], which will be stated in Proposition 7.5, we will assume that $n > 1700(\log n)^3$. Otherwise, we have $n \leq 1700(\log n)^3$ and therefore in this case $n \ll s$, and the previously established bound $\mathcal{O}(n)$ for the number of primitive solutions of general Thue's inequalities (see [1], for example) will prove Proposition 6.1.

Our goal now is to show (65), for $n > 1700(\log n)^3$. Our discussion is the same as Section 9 of [18].

Proposition 7.5. *Let $f(z)$ be a polynomial of degree n with rational coefficients, of Mahler height $M(f)$ and without multiple roots. Suppose that $f(x)f'(x)$ has not more than $q - 1$ real roots, where $q \geq 1$. Suppose further that $M(f) > e^{2n}$. Then for ϕ in*

$$(66) \quad 1700n^{-1}(\log n)^3 \leq \phi \leq 1,$$

the number of roots $\mathfrak{r} + \mathbf{i}\mathfrak{y}$ with imaginary part in $0 < \mathfrak{y} \leq M(f)^{-\phi}$ does not exceed $\left(\frac{8nq}{\phi}\right)^{1/2}$.

Proof. This is the Corollary in Section 9 of [18]. \square

Since $F \in C(4s - 2)$, then $f(z) = F(z, 1)$ has no more than $4s - 1$ real zeros and $f'(z) = F_x(z, 1)$ has no more than $4s - 2$ real zeros. So ff' has at most $8s - 3 = q - 1$ real zeros, where we take

$$(67) \quad q = 8s - 2.$$

We may suppose that $\Phi_1 \geq \dots \geq \Phi_n$. The number of summands in

$$(68) \quad \sum_{i=1}^n \min(1, \Phi_i)$$

with $\Phi_i \geq 1$ is the number of roots α_i of $F(x, 1)$ with $|\operatorname{Im} \alpha_i| \leq M^{-1}$. By taking $\phi = 1$ in (66), the contribution of summands with $\Phi_i \geq 1$ in the sum (68) is

$$< (8nq)^{1/2} = 2\sqrt{2nq}.$$

Clearly the summands with $\Phi_i \leq 1700n^{-1}(\log n)^3 < 1$ do not contribute.

The remaining summands have $1700n^{-1}(\log n)^3 < \Phi_i < 1$. Since $|\operatorname{Im} \alpha_j| = M^{-\Phi_j} \leq M^{-\Phi_i}$ for $j \leq i$, Proposition 7.5 yields $i < (8nq/\Phi_i)^{1/2}$, so $\Phi_i < 8nq/i^2$. We conclude that these terms contribute

$$\begin{aligned} &< \sum_{i=1}^n \min(1, 8nq/i^2) = [2\sqrt{2nq}] + 8nq \sum_{i=\lfloor \sqrt{8nq} \rfloor + 1}^n 1/i^2 \\ &< 4\sqrt{2nq} + 1 - 8q \ll \sqrt{ns}, \end{aligned}$$

by (67).

ACKNOWLEDGEMENTS

The authors are very grateful to the anonymous referee for very helpful suggestions. This article was written while Akhtari was a visitor at the *Max Planck Institute for Mathematics in Bonn*. Akhtari acknowledges the MPIM constant support for her research and collaboration, and in particular both authors are grateful for the opportunity to meet and advance this project in Bonn. The authors acknowledge the support from FIM, *Forschungsinstitut für Mathematik*, and are thankful to FIM for the hospitality and providing stimulating research atmosphere, especially during Akhtari's visits to Zürich in 2018 and 2019. Akhtari's research is partially supported by the NSF grant DMS-1601837 and *Simons Foundation's collaboration grant for mathematicians*. Bengoechea's research is supported by SNF grant 173976.

REFERENCES

- [1] S. Akhtari, Representation of small integers by binary forms, *Quart. J. Math. Oxford J.* 66 no. 4 (2015), 1009-1054.
- [2] S. Akhtari, N. Saradha, and D. Sharma, Thue's inequalities and the hypergeometric method, *the Ramanujan Journal* (2018) 45: 521-567.
- [3] M.A. Bennett, Rational approximation to algebraic numbers of small height: the Diophantine equation $|ax^n - by^n| = 1$. *J. Reine Angew. Math.* 535 (2001), 1-49.
- [4] B.J. Birch and J.R. Merriman, Finiteness theorems for binary forms with given discriminant, *Proc. London Math. Soc.* 25 (1972), 385-394.
- [5] E. Bombieri, W. M. Schmidt, On Thue's equation, *Invent. Math.* 88 (1987), 69-81.
- [6] J. H. Evertse, On the equation $ax^n - by^n = c$, *Compositio Math.* 47 (1982), 289-315.
- [7] J. H. Evertse and K. Györy, Thue inequalities with a small number of solutions, in: *The mathematical heritage of C.F. Gauss*, World Scientific Publ. Co., Singapore, 1991, 204-224.
- [8] J. H. Evertse and K. Györy, Effective finiteness results for binary forms with given discriminant, *Compositio Mathematica*, 79 (1991) no. 2, 169-204.
- [9] K. Györy, Thue inequalities with a small number of primitive solutions, *Period. Math. Hungar.* 42 (2001), no. 1-2, 199-209.

- [10] D. Lewis and K. Mahler, Representation of integers by binary forms, *Acta Arith.* 6 (1961), 333-363.
- [11] K. Mahler, Zur Approximation algebraischer Zahlen III. Über die mittlere Anzahl der Darstellungen grosser Zahlen durch binäre Formen, *Acta. Math.* 62 (1933), 91-166.
- [12] K. Mahler, On two extremum properties of polynomials, *Illinois J. Math.* 7 1(963) 681-701.
- [13] K. Mahler, An inequality for the discriminant of a polynomial, *Michigan Math. J.* 11 (1964), 257-262.
- [14] M. Mignotte, A note on the equation $ax^n - by^n = c$, *Acta Arith.* 75 (1996) no. 3, 287-295.
- [15] J. Mueller, Counting solutions of $|ax^r - by^r| \leq h$, *Quart. J. Math. Oxford Ser. (2)* 38 (1987), no. 152, 503-513.
- [16] J. Mueller and W. M. Schmidt, Trinomial Thue equations and inequalities, *J. Reine Angew. Math.*, 379 (1987), 76-99.
- [17] J. Mueller and W. M. Schmidt, Thue's equation and a conjecture of Siegel, *Acta Math.* 160 (1988), no. 3-4, 207-247.
- [18] W. M. Schmidt, Thue equations with few coefficients, *Trans. Amer. Math. Soc.* 303 (1987), 241-255.
- [19] W. M. Schmidt, *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Mathematics 1467, Springer Verlag, 1991.
- [20] C. L. Siegel, Die Gleichung $ax^n - by^n = c$, *Math. Ann.* 114 (1937), 57-68.
- [21] C. L. Stewart, On the number of solutions of polynomial congruences and Thue equations, *J. Amer. Math. Soc.* 4 (1991), 793-835.
- [22] A. Thue, Berechnung aller Lösungen gewisser Gleichungen von der form $ax^r - by^r = f$. *Vid. Skrifter I Mat.-Naturv. Klasse* (1918), 1-9.
- [23] J. L. Thunder, on Thue inequalities and a conjecture of Schmidt, *J. Number Theory* 52 (1995), 319-328.
- [24] J. L. Thunder, Decomposable form inequalities, *Ann. of Math. (2)* 153 (2001), no. 3, 767-804.

DEPARTMENT OF MATHEMATICS, FENTON HALL, UNIVERSITY OF OREGON, EUGENE,
OR 97403-1222 USA

E-mail address: akhtari@uoregon.edu

ETH, MATHEMATICS DEPT., CH-8092, ZÜRICH, SWITZERLAND

E-mail address: paloma.bengoechea@math.ethz.ch