ON WEYL'S INEQUALITY, HUA'S LEMMA, AND EXPONENTIAL SUMS OVER BINARY FORMS

Trevor D. Wooley*

1. Introduction

The remarkable success enjoyed by the Hardy-Littlewood method in its application to diagonal diophantine problems rests in large part on the theory of exponential sums in a single variable. Following almost a century of intense investigations, the latter body of knowledge has reached a mature state which, although falling short of what is expected to be true, nonetheless suffices for the majority of applications. By contrast, exponential sums in many variables remain poorly understood, and consequently applications of the Hardy-Littlewood method to problems concerning the solubility of systems of forms in many variables are fraught with difficulties. While the methods of Weyl and of Vinogradov have been extended to estimate exponential sums in many variables (see, in particular, Arkhipov and Karatsuba [1], Arkhipov, Karatsuba and Chubarikov [2], Tartakovsky [26], Davenport [11], [12], [13], [14], Birch [4] and Schmidt [25]), in almost all circumstances the strength of the ensuing estimates is considerably inferior to that of corresponding estimates for Weyl sums in a single variable, and moreover one is obliged to work under various hypotheses of a geometric nature. Indeed, it is only for exponential sums over polynomials diagonalisable over $\mathbb{C}$, and non-singular cubic polynomials and their close kin, that we have estimates of strength to match those for corresponding exponential sums of a single variable (see Chowla and Davenport [10], Birch and Davenport [5], Heath-Brown [20] and Hooley [21], [22], [23]). The purpose of this paper is to develop estimates for exponential sums over binary forms of strength comparable to the best available for corresponding exponential sums of a single variable, and, moreover, without any serious geometric hypotheses on the form. Since binary forms of degree exceeding 3 in general fail to diagonalise over $\mathbb{C}$, it should be evident that our conclusions go beyond those of Birch and Davenport [5]. Our hope is that the methods described herein may spawn ideas for improved treatments of exponential sums in many variables.

Before describing our main conclusions we require some notation. Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ is a binary form of degree d exceeding 1. Then we say that Φ is *degenerate* if there exist complex numbers α and β such that $\Phi(x,y)$ is identically equal to $(\alpha x+\beta y)^d$. It is easily verified that when $\Phi(x,y)$ is degenerate, then there exist integers a, b and c with $\Phi(x,y)=a(bx+cy)^d$. Our first theorem, which we establish in §3, provides an analogue of Weyl's inequality for exponential sums over non-degenerate binary forms. Throughout, we write $e(z)$ for $e^{2\pi iz}$.

Theorem 1. *Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ is a non-degenerate form of degree $d\geqslant 3$. Let $\alpha\in\mathbb{R}$, and suppose that there exist $r\in\mathbb{Z}$ and $q\in\mathbb{N}$ with $(r,q)=1$ and $|\alpha-r/q|\leqslant q^{-2}$. Finally, suppose that P and Q are real numbers sufficiently large in terms of the coefficients of Φ, and satisfying $P\asymp Q$. Then for each $\varepsilon>0$, one has*

$$\sum_{1\leqslant x\leqslant P}\sum_{1\leqslant y\leqslant Q}e(\alpha\Phi(x,y))\ll P^{2+\varepsilon}(q^{-1}+P^{-1}+qP^{-d})^{2^{2-d}}.$$

We remark that the conclusion of Theorem 1 is identical with that following from the classical version of Weyl's inequality (see, for example, Lemma 2.4 of Vaughan [27]) in circumstances where $\Phi(x,y)$ is a diagonal form. Moreover Theorem 1 of Chowla and Davenport [10] establishes the same conclusion as

---

2000 *Mathematics Subject Classification.* 11D72, 11L07, 11E76, 11P55, 11J71.

Key words and phrases. Exponential sums, binary forms, diophantine equations, distribution modulo 1.

*Packard Fellow and supported in part by NSF grant DMS-9622773. This paper was written while the author was enjoying the hospitality of the Department of Mathematics at Princeton University.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Theorem 1 above in the special case where $\Phi(x,y)$ is a binary cubic form with non-zero discriminant. In cases where $d>3$, meanwhile, the conclusion of Theorem 1 provides substantially sharper estimates than would be available through the work of Birch [4] and Schmidt [25]. When d is larger than 12 or so, a trivial argument employing Vinogradov's methods yields an estimate superior to that provided by Theorem 1. We discuss such estimates briefly in §8 below.

The differencing process which leads to the estimate recorded in Theorem 1 may be employed in a familiar manner to establish mean value estimates for exponential sums over binary forms analogous to Hua's Lemma. The details of such a treatment, unfortunately, require knowledge concerning the number of integral points on certain affine plane curves. Our current state of knowledge on such topics being incomplete, our conclusions are somewhat weaker than would otherwise be the case. Following some preliminaries in §§4 and 5, we establish our analogue of Hua's Lemma in §§6 and 7.

Theorem 2. *Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ is a non-degenerate form of degree $d\geqslant 3$. When $d=3$ or 4 and j is an integer with $1\leqslant j\leqslant d$, or when $d\geqslant 5$ and $j=1$ or 2, one has for each positive number ε the bound*

$$\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P}e(\alpha\Phi(x,y))\Big|^{2^{j-1}}d\alpha\ll P^{2^j-j+\varepsilon}.$$

When $d\geqslant 5$ and j is an integer with $1\leqslant j\leqslant d-1$, then for each positive number ε one has

$$\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P}e(\alpha\Phi(x,y))\Big|^{2^{j-1}}d\alpha\ll P^{2^j-j+\frac{1}{2}+\varepsilon}.$$

Finally, when $d\geqslant 5$, for each positive number ε one has

$$\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P}e(\alpha\Phi(x,y))\Big|^{\frac{5}{16}2^d}d\alpha\ll P^{\frac{5}{8}2^d-d+1+\varepsilon}$$

and

$$\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P}e(\alpha\Phi(x,y))\Big|^{\frac{9}{16}2^d}d\alpha\ll P^{\frac{9}{8}2^d-d+\varepsilon}.$$

In circumstances where $\Phi(x,y)$ is a diagonal form, the conclusion of Theorem 2 is identical with the classical version of Hua's Lemma for $d=3,4$ (see, for example, Lemma 2.5 of Vaughan [27]), and a little weaker by comparison with this diagonal situation when $d\geqslant 5$. Whereas Theorem 2 requires $\frac{9}{16}2^d$ copies of the generating function in order to obtain an optimal mean value estimate, the classical version of Hua's Lemma for diagonal forms would require only 2^{d-1}. This discrepency would be remedied by improved knowledge concerning the number of integer points on certain affine plane curves. Indeed, by rather involved considerations of absolute irreducibility criteria based on a division into many cases, it is possible to improve certain of the above exponents to a small extent. Such arguments being rather lengthy and technical, and in any case limited in their application, we have chosen to present the main thrust of our ideas and defer any such considerations to a possible future occasion. We note that when $d=3$ and the underlying form has non-zero discriminant, then one may establish the main conclusion of Theorem 2 through the methods of Chowla and Davenport [10] (see Lemma 4.1 of Brüdern and Wooley [9]). Finally, when d is larger than 11 or so, it is possible to apply a trivial argument involving Vinogradov's methods in order to obtain conclusions superior to those stemming from Theorem 2. We discuss such estimates in §8 below.

As our first application of the new estimates provided by Theorems 1 and 2, in §9 we consider the solubility of homogeneous diophantine equations which split as sums of binary forms. In order to contain our deliberations within this paper, we illustrate our ideas with a relatively simple conclusion typical of the kind attainable through the methods of §9.

Theorem 3. *Let d be an integer with $d\geqslant 3$, and define $s_0(d)$ by*

$$s_0(d)=\begin{cases}2^{d-1}, & \text{when } d=3,4,\\ \frac{9}{16}2^d, & \text{when } d\geqslant 5.\end{cases}$$

Let $s > s_0(d)$, and let $\Phi_j \in \mathbb{Z}[x,y]$ $(1\leqslant j\leqslant s)$ be homogeneous forms of degree d with non-zero discriminants. Let $\mathcal{N}(P) = \mathcal{N}_s(P;\mathbf{\Phi})$ denote the number of solutions of the diophantine equation

$$\Phi_1(x_1,y_1)+\cdots+\Phi_s(x_s,y_s)=0, \tag{1.1}$$

subject to $|x_j|\leqslant P$ and $|y_j|\leqslant P$ $(1\leqslant j\leqslant s)$. Then provided that the form $\Phi_1(x_1,y_1)+\cdots+\Phi_s(x_s,y_s)$ is indefinite, one has

$$\mathcal{N}_s(P;\mathbf{\Phi}) = \mathcal{C}\mathfrak{S}P^{2s-d} + O_{\mathbf{\Phi}}(P^{2s-d-\delta}),$$

for some positive number δ, where $\mathcal{C}$ denotes the volume of the $(2s-1)$-dimensional hypersurface determined by the equation (1.1) contained in the box $[-1,1]^s$, and $\mathfrak{S}$ denotes the singular series $\prod_p v_p$, where the product is over prime numbers, and

$$v_p = \lim_{h\to\infty} p^{h(1-2s)} M_s(p^h;\mathbf{\Phi}),$$

in which $M_s(p^h;\mathbf{\Phi})$ denotes the number of solutions of the congruence

$$\Phi_1(x_1,y_1)+\cdots+\Phi_s(x_s,y_s)\equiv 0 \pmod{p^h},$$

with $1\leqslant x_j, y_j\leqslant p^h$ $(1\leqslant j\leqslant s)$.

Corollary. *Under the hypotheses of Theorem 3, whenever*

$$s > \max\{s_0(d), d^2\},$$

and the form $\Phi_1(x_1,y_1)+\cdots+\Phi_s(x_s,y_s)$ is indefinite, one has

$$P^{2s-d} \ll_{\mathbf{\Phi}} \mathcal{N}_s(P;\mathbf{\Phi}) \ll_{\mathbf{\Phi}} P^{2s-d}.$$

An examination of the methods employed in §9 in the course of the proof of Theorem 3 will reveal that there is no difficulty in principle in obtaining an asymptotic formula for the number of solutions of the equation (1.1) satisfying $(\mathbf{x},\mathbf{y})\in \mathcal{D}P\cap\mathbb{Z}^{2s}$, for any convex subset $\mathcal{D}$ of $\mathbb{R}^{2s}$. Indeed, with extra effort the condition that the forms Φ_j have non-zero discriminant can also be removed, so long as the forms are at least non-degenerate. We have imposed the condition that the discriminants be non-zero in the statement of Theorem 3 in order to avoid discussion of possible singularities, but this is an entirely technical consideration. We remark that for larger d the permissible $s_0(d)$ may be reduced in line with the discussion of §8 below. In particular, when d is large, the main conclusion of the corollary to Theorem 3 holds with $s = d^2(\log d + \log\log d + O(1))$. Further, we note that the condition in the corollary that s exceed d^2 is imposed purely to guarantee the existence of non-singular p-adic solutions of the equation (1.1) for each prime p. We note that $s_0(d)\geqslant d^2$ for $d\geqslant 6$, and moreover that results on additive quintics in the literature enable the hypothesis of the corollary to be relaxed when $d=5$ to the condition that $s>18$. Presumably the expression d^2 in the statement of the corollary can be replaced by $\frac{1}{2}d^2$ (at least for odd d), but at this time there appears to be no work in the literature concerning the local solubility of equations of the shape (1.1) for $d\geqslant 4$. When $d=3$, of course, it follows from Theorem 1.1 of Brüdern and Wooley [9] that for each positive number ε, one has

$$P^{5-\varepsilon} \ll_{\varepsilon,\mathbf{\Phi}} \mathcal{N}_4(P;\mathbf{\Phi}) \ll_{\varepsilon,\mathbf{\Phi}} P^{5+\varepsilon}.$$

Finally, we note that the techniques discussed in §9 may be applied also to establish that all large integers n satisfying the necessary local solubility conditions are represented non-trivially by a sum of s non-degenerate binary forms of degree d, provided only that $s > s_0(d)$. Thus we make a non-trivial contribution to Problem 31 of Lewis [24].

A second application of Theorem 1, which we discuss in §10, concerns small values of binary forms modulo one. When $\theta\in\mathbb{R}$, we write $\|\theta\|$ for $\min_{y\in\mathbb{Z}}|\theta-y|$.

Theorem 4. *Let d be an integer with $d\geqslant 3$, and let $\Phi(x,y)\in\mathbb{Z}[x,y]$ be a non-degenerate homogeneous form of degree d. Suppose that α is a real number, and δ is any positive number. Then there is a real number $N(\delta,d)$ such that whenever $N\geqslant N(\delta,d)$, one has*

$$\min_{1\leqslant m,n\leqslant N}\|\alpha\Phi(m,n)\|\leqslant N^{\delta-2^{2-d}}.$$

The conclusion of the theorem is of the same strength as would be attained via classical methods (see Baker [3]) for diagonal forms. We remark that by fixing one of the variables in the binary form to be a multiple of the other, the methods of Wooley [32] yield sharper conclusions than those of Theorem 4 when d exceeds 8 or thereabouts.

It transpires that unpleasant subcases must be negotiated in our proofs of Theorems 1 and 2, and the ensuing complications obscure the simple ideas motivating our proofs. It may therefore be helpful to sketch the argument underlying the simplest case considered in Theorem 1. Consider then a binary cubic form with integral coefficients, say $\Phi(X,Y)=AX^3+BX^2Y+CXY^2+DY^3$. Our first idea is to "complete the cube" to obtain

$$\begin{aligned}27A^2\Phi(X,Y)=&(3AX+BY)^3+(9AC-3B^2)(3AX+BY)Y^2\\&+(27A^2D-9ABC+2B^3)Y^3.\end{aligned}$$

Thus, when A is non-zero, one may make a non-singular linear change of variables to convert the original form $\Phi(X,Y)$ into an associated form $\Psi(x,y)$ of the shape $\Psi(x,y)=ax^3+bxy^2+cy^3$, with $a\neq 0$. It transpires that such a transformation is possible, in fact, whenever $\Phi(X,Y)$ is non-degenerate, and indeed one then has that one at least of b and c is non-zero. Moreover, an exponential sum over $\Phi(X,Y)$ may be estimated in terms of an associated exponential sum over the polynomial $\Psi(x,y)$. We discuss such transformations in detail in §2. For the purposes of exposition, therefore, it suffices to estimate the exponential sum

$$h(\alpha)=\sum_{|x|\leqslant X}\sum_{|y|\leqslant X}e(\alpha(ax^3+bxy^2+cy^3)).$$

Since the situation in which $b=0$ is the diagonal case accessible to classical methods, we suppose that $b\neq 0$. Our second idea is to use an operation more trivial than Weyl differencing, that is, we apply Cauchy's inequality. Thus we obtain

$$\begin{aligned}|h(\alpha)|^2&\ll X\sum_{|x|\leqslant X}\Big|\sum_{|y|\leqslant X}e(\alpha(ax^3+bxy^2+cy^3))\Big|^2\\&=X\sum_{|x|\leqslant X}\sum_{|y_1|\leqslant X}\sum_{|y_2|\leqslant X}e(\alpha(bx(y_1^2-y_2^2)+c(y_1^3-y_2^3))).\end{aligned}\tag{1.2}$$

On interchanging the order of summation and employing a simple estimate for the divisor function, one obtains

$$\begin{aligned}|h(\alpha)|^2&\ll X\sum_{|y_1|\leqslant X}\sum_{|y_2|\leqslant X}\Big|\sum_{|x|\leqslant X}e(\alpha bx(y_1^2-y_2^2))\Big|\\&\ll X^3+X^{1+\varepsilon}\sum_{0<|m|\leqslant|b|X^2}\min\{X,\|m\alpha\|^{-1}\},\end{aligned}\tag{1.3}$$

and by comparison with the familiar situation of the cubic Weyl sum subject to two Weyl differencing steps, one obtains the desired analogue of Weyl's inequality recorded in Theorem 1. For forms of higher degree, one may proceed similarly to "complete the dth power", and then apply Cauchy's inequality. In a manner analogous to the process visible in (1.2), the corresponding exponential sum has degree at most $d-2$ with respect to one of the underlying variables. When that degree is precisely $d-2$, one may Weyl difference $d-3$ times in order to arrive at a familiar situation analogous to (1.3). If this degree is less than $d-2$, however, a more sophisticated approach is required. In any case, the total number of differencing steps, including the initial "trivial step", is only $d-2$, compared to $d-1$ in the classical treatment, and it is this advantage which is responsible for the relative success of our approach.

Throughout this paper, implicit constants occurring in Vinogradov's notation $\ll$ and $\gg$ will depend at most on the coefficients of the implicit binary forms, a small positive number ε, exponents d and k, and quantities occurring as subscripts to the latter notations, unless otherwise indicated. We write $f \asymp g$ when $f \ll g$ and $g \ll f$. When x is a real number, we write $[x]$ for the greatest integer not exceeding x. Also, we use vector notation for brevity. Thus, for example, the s-tuple $(\Phi_1, \dots, \Phi_s)$ will be abbreviated simply to $\mathbf{\Phi}$. In an effort to simplify our exposition, we adopt the convention that whenever ε appears in a statement, we are implicitly asserting that the statement holds for each $\varepsilon > 0$. Note that the "value" of ε may consequently change from statement to statement.

The author is grateful to the referee for useful comments.

2. Initial transformations

Let k be an integer with $k \geqslant 3$ and let $\Phi(x, y) \in \mathbb{Z}[x, y]$ be a non-degenerate homogeneous polynomial of degree k. Let P and Q be large real numbers with $P \asymp Q$, and define the exponential sum $F(\alpha) = F(\alpha; P, Q)$ by

$$F(\alpha; P, Q) = \sum_{1 \leqslant x \leqslant P} \sum_{1 \leqslant y \leqslant Q} e(\alpha \Phi(x, y)). \tag{2.1}$$

Before advancing to establish our main conclusions we transform the exponential sum $F(\alpha)$ into a related sum susceptible to our differencing procedure. We begin our discussion with a preliminary observation which permits us to save effort in our subsequent deliberations. Throughout, we abbreviate $\frac{\partial}{\partial x}\Phi(x, y)$ to $\Phi_x(x, y)$ and likewise $\frac{\partial}{\partial y}\Phi(x, y)$ to $\Phi_y(x, y)$.

Lemma 2.1. *Suppose that $\Theta(x, y) \in \mathbb{Z}[x, y]$ is a non-degenerate binary form of degree k exceeding 2. Then there exist integers a, b, c, d with the property that*

(i) $\Theta(a, c) \neq 0$,
(ii) $b\Theta_x(a, c) + d\Theta_y(a, c) = 0$,
(iii) $ad - bc \neq 0$,
(iv) $\Theta(ax + by, cx + dy)$ *depends explicitly on y.*

Proof. Observe that since $\Theta(x, y)$ is non-degenerate, then a trivial counting argument shows that there exist integers a and c, not both zero, with $\Theta(a, c) \neq 0$. This shows that the property (i) holds. Next, since the homogeneity of $\Theta(x, y)$ ensures that

$$x\Theta_x(x, y) + y\Theta_y(x, y) = k\Theta(x, y),$$

we find from (i) that

$$a\Theta_x(a, c) + c\Theta_y(a, c) \neq 0. \tag{2.2}$$

In particular, therefore, at least one of $\Theta_x(a, c)$ and $\Theta_y(a, c)$ is non-zero. Consequently, there exist integers b and d, with (b, d) linearly independent of (a, c), satisfying the equation

$$b\Theta_x(a, c) + d\Theta_y(a, c) = 0.$$

This confirms property (ii). It follows from the linear independence of (a, c) and (b, d), moreover, that $ad - bc \neq 0$, whence property (iii) also holds. Finally, since it follows from property (iii) that the transformation $(x, y) \longrightarrow (ax + by, cx + dy)$ is non-singular, one may conclude from the non-degeneracy of $\Theta(x, y)$ that the polynomial $\Theta(ax + by, cx + dy)$ must depend explicitly on both x and y. This establishes property (iv), and concludes the proof of the lemma.

We now relate the exponential sum $F(\alpha)$ to a related, though simpler, exponential sum. In this context it is useful, when $\Phi(x, y) \in \mathbb{Z}[x, y]$, to describe the polynomial Ψ as being a *condensation* of Φ when the following condition $(\mathcal{C})$ is satisfied.

$(\mathcal{C})$ We have $\Psi(u, v) \in \mathbb{Z}[u, v]$, and the coefficients of Ψ depend at most on those of Φ. Further, the polynomial $\Psi(u, v)$ has the same degree as $\Phi(x, y)$, and takes the shape

$$\Psi(u, v) = Au^k + Bu^{k-t}v^t + \sum_{j=t+1}^{k} C_j u^{k-j} v^j, \tag{2.3}$$

with $AB \neq 0$ and $2 \leqslant t \leqslant k$.

Lemma 2.2. *There is a condensation Ψ of Φ, a positive integer D depending at most on the coefficients of Φ, and a positive real number X with $X \asymp P$, satisfying the property that for every real number α one has*

$$|F(\alpha;P,Q)| \ll (\log X)^2 \sup_{\beta,\gamma\in\mathbb{R}} |H(\alpha/D;\beta,\gamma;X)|,$$

where

$$H(\theta;\beta,\gamma;X) = \sum_{|u|\leqslant X}\sum_{|v|\leqslant X} e(\theta\Psi(u,v)+\beta u+\gamma v). \tag{2.4}$$

Proof. We begin by recalling the hypotheses concerning Φ from the opening paragraph of this section, and we apply Lemma 2.1 to conclude that there exist integers a,b,c,d satisfying the properties (i)–(iv) with $\Theta = \Phi$. Define the polynomial $\Psi(u,v)$ by

$$\Psi(u,v) = \Phi(au+bv,cu+dv). \tag{2.5}$$

Then we can write $\Psi(u,v)$ in the shape (2.3) for a suitable natural number t with $1\leqslant t\leqslant k$, and integers A, B and C_j $(t+1\leqslant j\leqslant k)$. By property (i) of Lemma 2.1 with $\Theta=\Phi$, it follows from (2.5) that $A\neq 0$. Moreover property (ii) ensures that the coefficient of $u^{k-1}v$ in $\Psi(u,v)$ vanishes, so that one may take $t\geqslant 2$ in (2.3). Furthermore, in view of property (iv) of Lemma 2.1, we may suppose that B is non-zero. Then we may conclude that the polynomial Ψ satisfies the condensation property $(\mathcal{C})$ above.

Next write $\Delta = ad-bc$, and observe that property (iii) of Lemma 2.1 for $\Theta=\Phi$ ensures that $\Delta\neq 0$. Plainly Δ depends at most on the coefficients of Φ. It follows from (2.1) and (2.5) that

$$F(\alpha;P,Q) = \sum_{1\leqslant x\leqslant P}\sum_{1\leqslant y\leqslant Q} e\left(\alpha\Psi\left(\frac{dx-by}{\Delta},\frac{ay-cx}{\Delta}\right)\right).$$

On breaking the ranges of summation into arithmetic progressions modulo $|\Delta|$, we therefore conclude that

$$F(\alpha;P,Q) = \sum_{r=1}^{|\Delta|}\sum_{s=1}^{|\Delta|}\sum_{z_-\leqslant z\leqslant z_+}\sum_{w_-\leqslant w\leqslant w_+} e(\theta\Psi(Z,W)), \tag{2.6}$$

where we write $\theta = \alpha/\Delta^k$, and

$$\begin{aligned} z_- &= \min\{(1-r)/\Delta,(P-r)/\Delta\}, \quad z_+ = \max\{(1-r)/\Delta,(P-r)/\Delta\},\\ w_- &= \min\{(1-s)/\Delta,(Q-s)/\Delta\}, \quad w_+ = \max\{(1-s)/\Delta,(Q-s)/\Delta\},\\ Z &= d(z\Delta+r)-b(w\Delta+s) \quad\text{and}\quad W = a(w\Delta+s)-c(z\Delta+r). \end{aligned} \tag{2.7}$$

On recalling (2.4), and writing

$$X = \max\{|d|P+|b|Q,|c|P+|a|Q\}, \tag{2.8}$$

it follows from (2.6) and (2.7) via orthogonality that

$$F(\alpha;P,Q) = \sum_{r=1}^{|\Delta|}\sum_{s=1}^{|\Delta|}\int_0^1\int_0^1 H(\theta;\beta,\gamma;X)\mathcal{M}(\Delta\beta,\Delta\gamma)d\beta d\gamma,$$

where

$$\mathcal{M}(\beta,\gamma) = \sum_{\substack{z_-\leqslant z\leqslant z_+\\ w_-\leqslant w\leqslant w_+}} e(\beta(bw-dz+(bs-dr)/\Delta)-\gamma(aw-cz+(as-cr)/\Delta)).$$

Consequently,

$$|F(\alpha;P,Q)| \ll M(P,Q)\sup_{\beta,\gamma\in\mathbb{R}}|H(\theta;\beta,\gamma;X)|, \tag{2.9}$$

where

$$M(P,Q) = \int_0^1\int_0^1 \min\{P,\|\Delta(c\gamma-d\beta)\|^{-1}\}\min\{Q,\|\Delta(b\beta-a\gamma)\|^{-1}\}d\beta d\gamma.$$

But by exploiting periodicity modulo 1, and making a change of variables, we obtain

$$M(P,Q) \ll \int_0^1\int_0^1 \min\{P,\|\lambda\|^{-1}\}\min\{Q,\|\mu\|^{-1}\}d\lambda d\mu \ll (\log P)(\log Q).$$

Thus the proof of the lemma follows immediately from (2.9).

We also require an analogue of Lemma 2.2 suitable for investigating mean value estimates. It transpires that this is almost trivial.

Lemma 2.3. *There is a condensation Ψ of Φ, and a positive real number X with $X \asymp P$, with the property that for every natural number s one has*

$$\int_0^1 |F(\alpha; P, Q)|^{2s} d\alpha \ll \int_0^1 |H(\alpha; X)|^{2s} d\alpha,$$

where we write

$$H(\theta; X) = \sum_{|u| \leqslant X} \sum_{|v| \leqslant X} e(\theta \Psi(u, v)).$$

Proof. As in the argument of the first paragraph of the proof of Lemma 2.2, it follows from Lemma 2.1 that there exist integers a, b, c, d satisfying the properties (i)-(iv) with $\Theta = \Phi$. As in the proof of Lemma 2.2, it follows that the polynomial $\Psi(u, v)$ defined by (2.5) is a condensation of Φ, and moreover that

$$\int_0^1 |F(\alpha; P, Q)|^{2s} d\alpha = \int_0^1 \Big| \sum_{1 \leqslant x \leqslant P} \sum_{1 \leqslant y \leqslant Q} e(\alpha \Psi(dx - by, ay - cx)/\Delta^k) \Big|^{2s} d\alpha,$$

where $\Delta = ad - bc$. On considering the underlying diophantine equation, therefore, we find that

$$\int_0^1 |F(\alpha; P, Q)|^{2s} d\alpha \ll \int_0^1 \Big| \sum_{1 \leqslant x \leqslant P} \sum_{1 \leqslant y \leqslant Q} e(\alpha \Psi(dx - by, ay - cx)) \Big|^{2s} d\alpha$$
$$\ll \int_0^1 |H(\alpha; X)|^{2s} d\alpha,$$

where X is given by (2.8). This completes the proof of the lemma.

3. Weyl's inequality for binary forms

Before establishing Theorem 1, we arm ourselves with a technical lemma which provides a minor elaboration on the conclusion of Vaughan [27, Lemma 2.2]. We provide a proof for the sake of completeness.

Lemma 3.1. *Suppose that X, Y and α are real numbers with $X \geqslant 1$ and $Y \geqslant 1$. Suppose also that D is a positive integer, and that $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $|\alpha - a/q| \leqslant q^{-2}$ and $(a, q) = 1$. Then*

$$\sum_{1 \leqslant x \leqslant X} \min\{Y, \|\alpha x/D\|^{-1}\} \ll XY(Dq^{-1} + Y^{-1} + Dq(XY)^{-1}) \log(2DqX).$$

Proof. Plainly, it suffices to estimate the sum

$$S = \sum_{1 \leqslant x \leqslant X} \min\{XYx^{-1}, \|\alpha x/D\|^{-1}\}.$$

But by breaking up the summation into arithmetic progressions modulo qD, we obtain

$$S \leqslant \sum_{0 \leqslant j \leqslant X/(qD)} \sum_{r=1}^{qD} \min\left\{ \frac{XY}{Dqj + r}, \left\| \frac{\alpha}{D}(Dqj + r) \right\|^{-1} \right\}. \tag{3.1}$$

For each j occurring in the first summation of (3.1), write $y_j = [\alpha j q^2 D]$ and $\theta = q^2 \alpha - qa$. Then

$$\frac{\alpha}{D}(Dqj + r) = \frac{y_j + ar}{qD} + \frac{\{\alpha j q^2 D\}}{qD} + \frac{\theta r}{q^2 D}.$$

When $j = 0$ and $1 \leqslant r \leqslant \frac{1}{2}q$, therefore,

$$\left\| \frac{\alpha}{D}(Dqj + r) \right\| \geqslant \left\| \frac{ar}{qD} \right\| - \frac{1}{2qD} \geqslant \frac{1}{2} \left\| \frac{ar}{qD} \right\|.$$

Otherwise, for each j there are at most $O(D)$ values of r for which the inequality

$$\left\| \frac{\alpha}{D}(Dqj+r) \right\| \geqslant \frac{1}{2} \left\| \frac{y_j+ar}{qD} \right\|$$

fails to hold. Moreover, one has $Dqj + r \gg q(j+1)$ in this case. Thus we deduce that

$$\begin{aligned} S &\ll \sum_{1\leqslant r\leqslant q/2} \left\| \frac{ar}{qD} \right\|^{-1} + \sum_{0\leqslant j\leqslant X/(qD)} \left(\frac{XYD}{q(j+1)} + \sum_{\substack{r=1\\ Dq\nmid (y_j+ar)}}^{qD} \left\| \frac{y_j+ar}{qD} \right\|^{-1} \right) \\ &\ll DXYq^{-1} \sum_{0\leqslant j\leqslant X} \frac{1}{j+1} + (a,D)\left(\frac{X}{qD}+1 \right) \sum_{1\leqslant h\leqslant qD} \frac{qD}{h(a,D)}, \end{aligned}$$

and the desired conclusion follows immediately.

We note that in principle one could break the summation implicit in Lemma 3.1 into arithmetic progressions modulo D, and then appeal to Vaughan [27, Lemma 2.2]. Unfortunately, however, one then becomes entangled with sums involving expressions of the shape $\|\alpha u + \beta\|^{-1}$. We have opted here to avoid such annoying complications, minor as they may be.

We recall the Weyl differencing lemma. Let Δ_j denote the jth iterate of the forward differencing operator, so that for any function ϕ of a real variable α, one has

$$\Delta_1(\phi(\alpha);\beta) = \phi(\alpha+\beta) - \phi(\alpha),$$

and when j is a natural number,

$$\Delta_{j+1}(\phi(\alpha);\beta_1,\dots,\beta_{j+1}) = \Delta_1(\Delta_j(\phi(\alpha);\beta_1,\dots,\beta_j);\beta_{j+1}).$$

We adopt the convention that $\Delta_0(\phi(\alpha);\boldsymbol{\beta}) = \phi(\alpha)$.

Lemma 3.2. *Let X be a positive real number, and let ϕ be an arbitrary arithmetical function. Write*

$$T(\phi) = \sum_{|x|\leqslant X} e(\phi(x)).$$

Then for each natural number j there exist intervals $I_i = I_i(\mathbf{h})$ $(1\leqslant i\leqslant j)$, possibly empty, satisfying

$$I_1(h_1) \subseteq [-X,X] \quad \text{and} \quad I_i(h_1,\dots,h_i) \subseteq I_{i-1}(h_1,\dots,h_{i-1}) \quad (2\leqslant i\leqslant j),$$

with the property that

$$|T(\phi)|^{2^j} \leqslant (4X+1)^{2^j-j-1} \sum_{|h_1|\leqslant 2X} \cdots \sum_{|h_j|\leqslant 2X} T_j,$$

and here we write

$$T_j = \sum_{x\in I_j\cap\mathbb{Z}} e(\Delta_j(\phi(x);h_1,\dots,h_j)).$$

Proof. This is a trivial variant of Lemma 2.3 of Vaughan [27].

We are now in a position to prove Theorem 1. Suppose that $\Phi(x,y) \in \mathbb{Z}[x,y]$ is a non-degenerate form of degree $k\geqslant 3$, and that $\Psi(u,v)$ is a polynomial satisfying condition $(\mathcal{C})$. Let D be a positive integer depending at most on the coefficients of Φ. We suppose that $\alpha \in \mathbb{R}$, $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $|\alpha - a/q| \leqslant q^{-2}$ and $(a,q) = 1$, and when X is large we aim to estimate the exponential sum $H(\alpha/D;\beta,\gamma;X)$ defined by (2.4).

We start by considering the situation in which $\Psi(u,v)$ takes the shape (2.3) with $t = k$, so that for fixed integers A and B depending at most on the coefficients of Φ, one has $\Psi(u,v) = Au^k + Bv^k$. Then by (2.4) we have

$$H(\alpha/D;\beta,\gamma;X) = f(A\alpha/D;\beta;X) f(B\alpha/D;\gamma;X), \tag{3.2}$$

where

$$f(\eta;\theta;X)=\sum_{|u|\leqslant X}e(\eta u^k+\theta u).$$

Write C for either A or B, and apply Lemma 3.2 with $j=k-1$. Then following a simple calculation (see, for example, the proof of Vaughan [27, Lemma 2.4]), one obtains for each real number θ the estimate

$$|f(C\alpha/D;\theta;X)|^{2^{k-1}}\ll X^{2^{k-1}-k}\sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{k-1}|\leqslant 2X}\sum_{x\in I_{k-1}}e\left(\frac{C}{D}h_1\dots h_{k-1}p_{k-1}(x;\mathbf{h})\right),\tag{3.3}$$

where I_{k-1} is an interval of integers contained in $[-X,X]$, and

$$p_{k-1}(x;\mathbf{h})=\tfrac{1}{2}k!\alpha(2x+h_1+\dots+h_{k-1}).$$

The number of terms counted by the summation in (3.3) with $h_1\dots h_{k-1}$ equal to zero is $O(X^{k-1})$. Thus, on applying a familiar bound to estimate the sum over x in (3.3), and making use of a simple estimate for the divisor function, we obtain

$$|f(C\alpha/D;\theta;X)|^{2^{k-1}}\ll X^{2^{k-1}-k}\Big(X^{k-1}+X^{\varepsilon}\sum_{1\leqslant h\leqslant H}\min\{X,\|h\alpha/D\|^{-1}\}\Big),\tag{3.4}$$

where $H=|C|k!(2X)^{k-1}$. Plainly, when $q\geqslant X^k$ one has $|f(C\alpha/D;\theta;X)|\leqslant 2X+1$ via a trivial estimate. When $1\leqslant q<X^k$, meanwhile, on recalling that D depends at most on the coefficients of Φ, it follows from (3.4) and Lemma 3.1 that

$$|f(C\alpha/D;\theta;X)|\ll X^{1+\varepsilon}(q^{-1}+X^{-1}+qX^{-k})^{2^{1-k}}.\tag{3.5}$$

Thus the estimate (3.5) in fact holds no matter how large q may be. We may therefore conclude from (3.2) that

$$|H(\alpha/D;\beta,\gamma;X)|\ll X^{2+\varepsilon}(q^{-1}+X^{-1}+qX^{-k})^{2^{2-k}}.\tag{3.6}$$

Consider next the situation in which $\Psi(u,v)$ takes the shape (2.3) with $2\leqslant t\leqslant k-1$. Then for fixed integers C_j $(t\leqslant j\leqslant k)$ depending at most on the coefficients of Φ, one has

$$\Psi(u,v_1)-\Psi(u,v_2)=\sum_{j=t}^{k}C_ju^{k-j}(v_1^j-v_2^j).$$

Write

$$J_\gamma(\theta;u;X)=\sum_{|v|\leqslant X}e\Big(\theta\sum_{j=t}^{k}C_ju^{k-j}v^j+\gamma v\Big).$$

Then following an application of Cauchy's inequality to (2.4), we obtain

$$|H(\theta;\beta,\gamma;X)|^2\leqslant(2X+1)\sum_{|u|\leqslant X}|J_\gamma(\theta;u;X)|^2.\tag{3.7}$$

But Lemma 3.2 yields

$$|J_\gamma(\theta;u;X)|^{2^{t-1}}\ll X^{2^{t-1}-t}\sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{t-1}|\leqslant 2X}\sum_{x\in I_{t-1}}e(p_\gamma(x;\theta;u;\mathbf{h})),\tag{3.8}$$

where I_{t-1} is an interval of integers contained in $[-X,X]$, and the polynomial $p_\gamma(x;\theta;u;\mathbf{h})$ is defined by

$$p_\gamma(x;\theta;u;\mathbf{h})=\theta\sum_{j=t}^{k}C_ju^{k-j}\Delta_{t-1}(x^j;h_1,\dots,h_{t-1})+\gamma\Delta_{t-1}(x;h_1,\dots,h_{t-1}).\tag{3.9}$$

We remark that the final term depending on γ in (3.9) will vanish whenever $t\geqslant 3$.

On applying Hölder's inequality to (3.7), we obtain

$$|H(\theta;\beta,\gamma;X)|^{2^{t-1}}\ll X^{2^{t-1}-1}\sum_{|u|\leqslant X}|J_\gamma(\theta;u;X)|^{2^{t-1}},$$

and hence we deduce from (3.8) and (3.9) that

$$|H(\theta;\beta,\gamma;X)|^{2^{t-1}}\ll X^{2^t-t-1}\sum_{|h_1|\leqslant 2X}\dots\sum_{|h_{t-1}|\leqslant 2X}\sum_{x\in I_{t-1}}|K(\theta;X;\mathbf{h};x)|,\tag{3.10}$$

where

$$K(\theta;X;\mathbf{h};x)=\sum_{|u|\leqslant X}e\Big(\theta\sum_{j=t}^k C_ju^{k-j}\Delta_{t-1}(x^j;h_1,\dots,h_{t-1})\Big).$$

But now Lemma 3.2 yields

$$\begin{aligned}|K(\theta;X;&\mathbf{h};x)|^{2^{k-t-1}}\\ &\ll X^{2^{k-t-1}-k+t}\sum_{|g_1|\leqslant 2X}\dots\sum_{|g_{k-t-1}|\leqslant 2X}\Big|\sum_{y\in I_{k-t-1}}e(\theta q(y;\mathbf{g},\mathbf{h};x))\Big|,\end{aligned}\tag{3.11}$$

where I_{k-t-1} is an interval of integers contained in $[-X,X]$, and the polynomial $q(y;\mathbf{g},\mathbf{h};x)$ is defined by

$$\begin{aligned}q(y;\mathbf{g},\mathbf{h};x)=&C_t\Delta_{k-t-1}(y^{k-t};\mathbf{g})\Delta_{t-1}(x^t;\mathbf{h})\\ &+C_{t+1}\Delta_{k-t-1}(y^{k-t-1};\mathbf{g})\Delta_{t-1}(x^{t+1};\mathbf{h}).\end{aligned}\tag{3.12}$$

Applying Hölder's inequality next to (3.10), we deduce from (3.11) and (3.12) that

$$\begin{aligned}|H(\theta;&\beta,\gamma;X)|^{2^{k-2}}\\ &\ll X^{2^{k-1}-2^{k-t-1}-t}\sum_{|h_1|\leqslant 2X}\dots\sum_{|h_{t-1}|\leqslant 2X}\sum_{x\in I_{t-1}}|K(\theta;X;\mathbf{h};x)|^{2^{k-t-1}}\\ &\ll X^{2^{k-1}-k}\sum_{|h_1|\leqslant 2X}\dots\sum_{|h_{t-1}|\leqslant 2X}\sum_{|g_1|\leqslant 2X}\dots\sum_{|g_{k-t-1}|\leqslant 2X}\sum_{x\in I_{t-1}}\Upsilon(\theta;\mathbf{g},\mathbf{h};x),\end{aligned}\tag{3.13}$$

where

$$\Upsilon(\theta;\mathbf{g},\mathbf{h};x)=\Big|\sum_{y\in I_{k-t-1}}e(\theta q(y;\mathbf{g},\mathbf{h};x))\Big|.\tag{3.14}$$

We next observe that following a simple calculation (see, for example, the proof of Vaughan [27, Lemma 2.4]), one has that $\Delta_{k-t-1}(y^{k-t-1};\mathbf{g})$ is independent of y, and further that

$$\Delta_{k-t-1}(y^{k-t};\mathbf{g})=\tfrac{1}{2}(k-t)!g_1\dots g_{k-t-1}(2y+g_1+\dots+g_{k-t-1})$$

and

$$\Delta_{t-1}(x^t;\mathbf{h})=\tfrac{1}{2}t!h_1\dots h_{t-1}(2x+h_1+\dots+h_{t-1}).$$

Consequently, when the expression

$$(2x+h_1+\dots+h_{t-1})h_1\dots h_{t-1}g_1\dots g_{k-t-1}\tag{3.15}$$

is non-zero, it follows from (3.12) and (3.14) that

$$\begin{aligned}\Upsilon(&\theta;\mathbf{g},\mathbf{h};x)\\ &\ll\min\{X,\|\tfrac{1}{2}C_t(k-t)!t!(2x+h_1+\dots+h_{t-1})h_1\dots h_{t-1}g_1\dots g_{k-t-1}\theta\|^{-1}\}.\end{aligned}$$

Since there are at most $O(X^{k-2})$ values of x, $\mathbf{g}$, $\mathbf{h}$ counted in the summations concluding (3.13) for which the expression (3.15) is zero, it follows from (3.13) via an elementary estimate for the divisor function that

$$|H(\alpha/D;\beta,\gamma;X)|^{2^{k-2}} \ll X^{2^{k-1}-k}\Big(X^{k-1}+X^{\varepsilon}\sum_{1\leqslant h\leqslant G}\min\{X,\|h\alpha/D\|^{-1}\}\Big), \tag{3.16}$$

where $G=t!(k-t)!|C_t|t(2X)^{k-1}$. Plainly, when $q\geqslant X^k$ one has

$$|H(\alpha/D;\beta,\gamma;X)|\leqslant(2X+1)^2$$

via a trivial estimate. When $1\leqslant q<X^k$, meanwhile, on recalling that D depends at most on the coefficients of Φ, it follows from (3.16) and Lemma 3.1 that the estimate (3.6) holds once again. Thus the estimate (3.6) holds no matter how large q may be.

Finally, we apply Lemma 2.2 to deduce in combination with (3.6) that whenever $\Phi(u,v)$ is non-degenerate, then there is a condensation Ψ of Φ, a positive integer D depending at most on the coefficients of Φ, and a positive real number X with $X\asymp P$, such that

$$\begin{aligned}|F(\alpha;P,Q)| &\ll (\log X)^2\sup_{\beta,\gamma\in\mathbb{R}}|H(\alpha/D;\beta,\gamma;X)|\\ &\ll X^{2+\varepsilon}(q^{-1}+X^{-1}+qX^{-k})^{2^{2-k}}.\end{aligned}$$

This completes the proof of Theorem 1.

4. Integral points on affine plane curves

The proof of our analogue of Hua's Lemma for exponential sums over binary forms depends for its success on estimates for the number of integral points on affine plane curves. In this section we record the necessary estimates for later use. Our basic tool is the following result of Bombieri and Pila [7] (we remark that earlier less precise conclusions would suffice for our purposes).

Lemma 4.1. *Let $\mathcal{C}$ be the curve defined by the equation $F(x,y)=0$, where $F(x,y)\in\mathbb{R}[x,y]$ is an absolutely irreducible polynomial of degree $d\geqslant 2$. Also, let $N\geqslant\exp(d^6)$. Then the number of integral points on $\mathcal{C}$, and inside a square $[0,N]\times[0,N]$, does not exceed*

$$N^{1/d}\exp(12(d\log N\log\log N)^{1/2}).$$

Proof. This is Theorem 5 of Bombieri and Pila [7].

In most applications of the above lemma, one is forced to enter into non-trivial discussions concerning the absolute irreducibility of polynomials, but in our present situation we are able largely to avoid such deliberations by averaging.

Lemma 4.2. *Let $\mathcal{N}$ denote a non-empty set of integers, associated to each one of which is a positive real number $w(n)$. Also, let X denote a large real number. Suppose that $F(x,y)\in\mathbb{Z}[x,y]$ is a non-degenerate polynomial of degree $d\geqslant 2$, and that X is sufficiently large in terms of d. Suppose also that for some fixed positive number A one has that the coefficients of F are each bounded in absolute value by X^A, and moreover that for each element n of $\mathcal{N}$ one has $|n|\leqslant X^A$. Let $E(X;\mathcal{N})$ denote the number of solutions of the diophantine equation $F(x,y)=n$, with $n\in\mathcal{N}$ and $(x,y)\in[-X,X]^2\cap\mathbb{Z}^2$, and with each solution (x,y,n) counted with weight $w(n)$. Then for each positive number ε, one has*

$$E(X;\mathcal{N})\ll X^{1/2+\varepsilon}\sum_{n\in\mathcal{N}}w(n)+X\max_{n\in\mathcal{N}}w(n),$$

where the implicit constant depends at most on d, ε and A, and otherwise is independent of the coefficients of F.

Proof. Consider some integer $n\in\mathcal{N}$. If $F(x,y)-n$ is absolutely irreducible, then the weighted number $\mathcal{B}(n)$ of solutions of the equation $F(x,y)=n$, with $(x,y)\in[-X,X]^2\cap\mathbb{Z}^2$, may be estimated by means of Lemma 4.1. Thus we obtain

$$\mathcal{B}(n)\ll X^{1/2+\varepsilon}w(n). \tag{4.1}$$

If $F(x,y)-n$ is not absolutely irreducible, then the non-degeneracy of $F(x,y)$ ensures that in this instance one has

$$\mathcal{B}(n)\ll Xw(n). \tag{4.2}$$

In this latter circumstance, moreover, one may write the polynomial $F(x,y)-n$ as a product of absolutely irreducible factors, say

$$F(x,y)-n=\prod_{j=1}^{l} g_j(x,y)\prod_{k=1}^{m} h_k(x,y), \tag{4.3}$$

where $l+m\geqslant 2$, and where $g_j(x,y)\in\mathbb{R}[x,y]$ $(1\leqslant j\leqslant l)$, and

$$h_k(x,y)=u_k(x,y)+v_k(x,y)\sqrt{-1}\quad (1\leqslant k\leqslant m),$$

with $u_k,v_k\in\mathbb{R}[x,y]$. Since $h_k(x,y)$ is presumed to be absolutely irreducible, we may suppose that $u_k(x,y)$ and $v_k(x,y)$ have no non-trivial polynomial common divisor over $\mathbb{C}[x,y]$. It therefore follows from Bezout's Theorem that the number of solutions of the simultaneous equations $u_k(x,y)=v_k(x,y)=0$ is bounded above by d^2. By considering real and imaginary components, therefore, the number of integral solutions of the equation $h_k(x,y)=0$ is also bounded above by d^2. If the degree of $g_j(x,y)$ exceeds 1 for any j, then the absolute irreducibility of $g_j(x,y)$ ensures, via Lemma 4.1, that the number of integral solutions of the equation $g_j(x,y)=0$, with $(x,y)\in[-X,X]^2\cap\mathbb{Z}^2$, is $O(X^{1/2+\varepsilon})$. Consequently, it follows from (4.3) that the estimate (4.1) can fail to hold only if some $g_j(x,y)$ is a linear polynomial.

Suppose then that for some j with $1\leqslant j\leqslant l$, the polynomial $g_j(x,y)$ is linear. If $g_j(x,y)$ is not some constant multiple of a $\mathbb{Q}$-rational linear polynomial, then since $g_j(x,y)$ is necessarily a constant multiple of a linear polynomial with algebraic coefficients, we deduce that the number of integral solutions of the equation $g_j(x,y)=0$ is at most one. For we may remove the constant factor and consider components with respect to some basis for the field extension containing the coefficients of $g_j(x,y)$. Then since $g_j(x,y)$ is not a constant multiple of a $\mathbb{Q}$-rational polynomial, it follows that the integral zeros of the equation $g_j(x,y)=0$ necessarily satisfy at least two linearly independent $\mathbb{Q}$-rational linear equations, whence the desired conclusion follows. Thus far we have shown that if the estimate (4.1) fails to hold, then necessarily the polynomial $F(x,y)-n$ is divisible by some $\mathbb{Q}$-rational linear polynomial. We next examine the set $\mathcal{N}^*$ of those $n\in\mathcal{N}$ for which the latter situation occurs.

Suppose next that n_0 is an integer satisfying the property that the polynomial $F(x,y)-n_0$ is divisible by the $\mathbb{Q}$-rational polynomial $ax+by+c$. There is plainly no loss of generality in supposing a, b and c to be integral, and we may choose integers a' and b' so that the polynomials $ax+by$ and $a'x+b'y$ are linearly independent. Write $\xi=ax+by+c$ and $\eta=a'x+b'y$. Then it follows that there is a polynomial $G(\xi,\eta)\in\mathbb{Z}[\xi,\eta]$, and a positive integer Δ, such that

$$\Delta(F(x,y)-n_0)=\xi G(\xi,\eta). \tag{4.4}$$

Before proceeding further we briefly investigate the dependence of $G(\xi,\eta)$ on η. Observe first that since $F(x,y)$ is non-degenerate, then necessarily $G(\xi,\eta)$ is explicit in η. Thus the polynomial $\frac{\partial}{\partial\eta}G(\xi,\eta)$ is non-trivial as a polynomial in η, and the set of its coefficients is a non-empty set consisting of certain non-trivial polynomials in ξ of degree at most $d-2$. Consequently, the set $\mathcal{S}$ of integers ξ_0 for which $G(\xi_0,\eta)$ is independent of η can have at most $d-2$ elements. We write $\mathcal{N}_0$ for the set of integers n for which

$$\Delta(n-n_0)=\xi_0 G(\xi_0,\eta),$$

for some $\xi_0\in\mathcal{S}$. Note, in particular, that $\text{card}(\mathcal{N}_0)\leqslant d-2$.

Let $n\in\mathcal{N}^*\setminus\mathcal{N}_0$, and suppose that $n\neq n_0$. Then it follows from (4.4) that the number of solutions of the diophantine equation $F(x,y)=n$ is bounded above by the number of integral solutions of the equation

$$\xi G(\xi,\eta)=\Delta(n-n_0), \tag{4.5}$$

with $\xi=ax+by+c$ and $\eta=a'x+b'y$ for some $(x,y)\in[-X,X]^2\cap\mathbb{Z}^2$. Since $n\neq n_0$, the number of divisors of $\Delta(n-n_0)$ is $O(X^\varepsilon)$. Thus there are at most $O(X^\varepsilon)$ choices for ξ satisfying (4.5). Let ξ_1 be any one such, and consider the equation

$$G(\xi_1,\eta)=\Delta(n-n_0)/\xi_1. \tag{4.6}$$

On noting that $n \notin \mathcal{N}_0$, we find that $\xi_1 \notin \mathcal{S}$, and hence that $G(\xi_1, \eta)$ depends explicitly on η. Thus the number of solutions in η of (4.6) is at most $d-1$. Consequently, whenever $n \in \mathcal{N}^* \setminus (\mathcal{N}_0 \cup \{n_0\})$, one has

$$\mathcal{B}(n) \ll X^{\varepsilon} w(n). \tag{4.7}$$

When $n \in \mathcal{N}_0 \cup \{n_0\}$, meanwhile, one may apply the estimate (4.2). On recalling that $\mathrm{card}(\mathcal{N}_0) \leqslant d-2$, therefore, we conclude from (4.1), (4.2) and (4.7) that

$$\begin{aligned} E(X;\mathcal{N}) &= \sum_{n\in\mathcal{N}} \mathcal{B}(n) \\ &\ll X^{1/2+\varepsilon} \sum_{n\in\mathcal{N}\setminus\mathcal{N}^*} w(n) + X^{\varepsilon} \sum_{n\in\mathcal{N}^*\setminus(\mathcal{N}_0\cup\{n_0\})} w(n) + X \sum_{n\in\mathcal{N}_0\cup\{n_0\}} w(n) \\ &\ll X^{1/2+\varepsilon} \sum_{n\in\mathcal{N}} w(n) + X \max_{n\in\mathcal{N}} w(n). \end{aligned}$$

This completes the proof of the lemma.

The situation for quadratic equations is particularly easy to handle.

Lemma 4.3. *Let a, b, c be integers with $abc \neq 0$, and let $S(a, b, c; P)$ denote the number of integral solutions of the equation $ax^2 + by^2 = c$, with $|x| \leqslant P$ and $|y| \leqslant P$. Then for each positive number ε, one has $S(a, b, c; P) \ll 1 + (|abc|P)^{\varepsilon}$.*

Proof. This estimate is well-known (see, for example, Estermann [17] or Vaughan and Wooley [28, Lemma 3.5]).

5. Hua's Lemma for binary forms: preliminaries

The object of this and the next two sections is to establish the mean value estimates recorded in Theorem 2. We begin here with some simplifying observations. Let $\Phi(x, y) \in \mathbb{Z}[x, y]$ be a non-degenerate binary form of degree $k \geqslant 3$. In view of Lemma 2.3, in order to establish Theorem 2 it suffices to consider the situation in which $\Phi(x, y)$ takes the shape

$$\Phi(x, y) = Ax^k + Bx^{k-t}y^t + \sum_{j=t+1}^{k} C_j x^{k-j} y^j, \tag{5.1}$$

with $AB \neq 0$ and $2 \leqslant t \leqslant k$. When X is a large real number, define

$$H(\alpha; X) = \sum_{|x|\leqslant X} \sum_{|y|\leqslant X} e(\alpha\Phi(x, y)), \tag{5.2}$$

and write also

$$F(\alpha; X) = \sum_{|u|\leqslant X} e(\alpha u^k). \tag{5.3}$$

Also, when j is a natural number, write

$$I_j(X) = \int_0^1 |H(\alpha; X)|^{2^{j-1}} d\alpha. \tag{5.4}$$

We establish our most general conclusions by induction, using the following mean value estimate as a starting point.

Lemma 5.1. *When $\Phi(x, y) \in \mathbb{Z}[x, y]$ is a non-degenerate form of degree $k \geqslant 3$, and X is a large real number, one has for each positive number ε the bound*

$$I_j(X) \ll X^{2^j - j + \varepsilon} \quad (j = 1, 2).$$

Proof. In view of (5.4), the case $j = 1$ follows immediately from the case $j = 2$ through an application of Schwarz's inequality. It suffices, therefore, to estimate $I_2(X)$. But by orthogonality, it follows from (5.2) and (5.4) that

$$I_2(X) = M(X), \tag{5.5}$$

where $M(X)$ denotes the number of integral solutions of the equation

$$\Phi(x_1, y_1) = \Phi(x_2, y_2), \tag{5.6}$$

with $|x_i|\leqslant X$ and $|y_i|\leqslant X$ $(i = 1, 2)$. For each integer h, write $r(h)$ for the number of representations of h in the form $\Phi(x, y) = h$, with $|x|\leqslant X$ and $|y|\leqslant X$. Since $\Phi(x, y)$ is non-degenerate, one plainly has $r(0) = O(X)$. We claim that when h is non-zero, one has $r(h) = O(X^\varepsilon)$, whence it follows from (5.6) that

$$\begin{aligned} M(X) = \sum_{h\in\mathbb{Z}} r(h)^2 &\leqslant r(0)^2 + (\max_{h\in\mathbb{Z}} r(h)) \sum_{h\in\mathbb{Z}} r(h) \\ &\ll X^2 + X^\varepsilon \sum_{h\in\mathbb{Z}} r(h) \ll X^{2+\varepsilon}. \end{aligned}$$

On recalling (5.5), we find that the conclusion of the lemma follows immediately from the latter bound.

In order to establish the above claim, we write $\Phi(x, y)$ as a product of irreducible factors over $\mathbb{Z}[x, y]$, say

$$\Phi(x, y) = \Psi_1(x, y)\Psi_2(x, y)\dots\Psi_r(x, y). \tag{5.7}$$

If h is a non-zero integer and $\Phi(x, y) = h$, then each $\Psi_i(x, y)$ is a divisor of h. Familiar estimates for the divisor function therefore show that there are at most $O(h^\varepsilon)$ possible choices for integers d_i $(1\leqslant i\leqslant r)$ such that $d_1d_2\dots d_r = h$ and $\Psi_i(x, y) = d_i$ $(1\leqslant i\leqslant r)$. But when a non-zero integer l has at most t distinct prime divisors, and the degree of Ψ_i exceeds 2, then it follows from Bombieri and Schmidt [8] that

$$\text{card}\{(x, y) \in \mathbb{Z}^2 \ :\ \Psi_i(x, y) = l\} \ll (\deg(\Psi_i))^{1+t} \ll l^\varepsilon. \tag{5.8}$$

Moreover, one may apply Lemma 4.3 to show that when Ψ_i has degree 2, one has

$$\text{card}\{(x, y) \in [-X, X]^2 \cap \mathbb{Z}^2 \ :\ \Psi_i(x, y) = l\} \ll (lX)^\varepsilon. \tag{5.9}$$

On recalling (5.7), therefore, and noting that $r(h)$ is non-zero only when $h = O(X^k)$, we conclude from (5.8) and (5.9) that

$$r(h) \ll (hX)^\varepsilon \ll X^{2\varepsilon}, \tag{5.10}$$

except possibly in circumstances where the decomposition (5.7) consists only of linear polynomials. Suppose that the latter is indeed the case, and further that two of these polynomials are linearly independent, say Ψ_l and Ψ_m. Since there can be at most one integral solution to the simultaneous equations $\Psi_l(x, y) = d_l$ and $\Psi_m(x, y) = d_m$, it again follows that the bound (5.10) holds. The only remaining case to consider is that in which the decomposition (5.7) satisfies the property that each $\Psi_i(x, y)$ is linear, and is a constant multiple of $\Psi_1(x, y)$. But then $\Phi(x, y) = \kappa(\Psi_1(x, y))^k$, for some real number κ, and thus $\Phi(x, y)$ is degenerate. This contradicts the hypotheses of the lemma, and so the proof of the lemma is complete.

Next we dispose swiftly of the diagonal case.

Lemma 5.2. *Suppose that $\Phi(x, y) \in \mathbb{Z}[x, y]$ is a non-degenerate binary form of degree $k\geqslant 2$ of the shape (5.1) with $t = k$. Let X be a large real number. Then for each integer j with $1\leqslant j\leqslant k$, and for each positive number ε, one has*

$$I_j(X) \ll X^{2^j - j + \varepsilon}.$$

Proof. In view of (5.1)-(5.3), when $t = k$ one may write

$$H(\alpha; X) = F(A\alpha; X)F(B\alpha; X),$$

and thus an application of Schwarz's inequality to (5.4), together with a change of variables, reveals that for each j with $1\leqslant j\leqslant k$, one has

$$\begin{aligned} I_j(X) &\ll \Big(\int_0^1 |F(A\alpha; X)|^{2^j}\, d\alpha\Big)^{1/2}\Big(\int_0^1 |F(B\alpha; X)|^{2^j}\, d\alpha\Big)^{1/2} \\ &\ll \int_0^1 |F(\alpha; X)|^{2^j}\, d\alpha. \end{aligned}$$

The desired conclusion therefore follows immediately from the classical version of Hua's Lemma (see, for example, Lemma 2.5 of Vaughan [27]).

Consider next the situation in which $\Phi(x,y)$ takes the shape (5.1) with $k\geqslant 4$ and $t=k-1$, so that for some integers a,b,c with $ab\ne 0$, one has

$$\Phi(x,y)=ax^k+bxy^{k-1}+cy^k. \tag{5.11}$$

If $c\ne 0$, we may write $u=kcy+bx$ and $v=x$, and thereby obtain an identity of the form

$$\Delta\Phi(x,y)=\Psi(u,v),$$

where Δ is some non-zero integer, and $\Psi(u,v)$ takes the shape (2.3) with $AB\ne 0$ and $t=2$. In such circumstances, on writing

$$G(\alpha;Y)=\sum_{|u|\leqslant Y}\sum_{|v|\leqslant Y}e(\alpha\Psi(u,v)),$$

it follows via the argument of the proof of Lemma 2.3 that for every natural number s, one has

$$\int_0^1|H(\alpha;X)|^{2s}d\alpha\ll\int_0^1|G(\alpha;Y)|^{2s}d\alpha,$$

where Y is a real number exceeding X by a factor depending only on the coefficients of $\Phi(x,y)$. We discuss the situations in which $k=3$, or else $k\geqslant 4$ and $2\leqslant t\leqslant k-2$, in §§6 and 7 below. Consequently, when $k\geqslant 4$ and $t=k-1$, we may restrict attention to polynomials of the shape (5.11) with $c=0$ and $ab\ne 0$.

Lemma 5.3. *Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ has the shape* (5.11) *with $c=0$ and $ab\ne 0$. Suppose also that $k\geqslant 3$ and that X is a large real number. Then for each integer j with $1\leqslant j\leqslant k$, and for each positive number ε, one has*

$$I_j(X)\ll X^{2^j-j+\varepsilon}. \tag{5.12}$$

Proof. The bound (5.12) is immediate from Lemma 5.1 when $j=1,2$. Suppose then that j is an integer with $2\leqslant j\leqslant k-1$, and that the inequality (5.12) holds. We aim to show that (5.12) holds also with j replaced by $j+1$, whence the upper bound (5.12) holds for each j satisfying $1\leqslant j\leqslant k$.

We first remove the contribution to $I_{j+1}(X)$ arising from a trivial part of the exponential sum $H(\alpha;X)$. Plainly,

$$|H(\alpha;X)|\ll X+\Big|\sum_{1\leqslant|x|\leqslant X}\sum_{|y|\leqslant X}e(\alpha(ax^k+bxy^{k-1}))\Big|.$$

Write

$$h(\alpha;X)=\sum_{|y|\leqslant X}e(\alpha y^{k-1}).$$

Then it follows from (5.4) via Hölder's inequality that

$$\begin{aligned}I_{j+1}(X)&\ll X^{2^{j-1}}I_j(X)+\int_0^1\Big|H(\alpha;X)\sum_{1\leqslant|x|\leqslant X}\sum_{|y|\leqslant X}e(\alpha(ax^k+bxy^{k-1}))\Big|^{2^{j-1}}d\alpha\\&\ll X^{2^{j-1}}I_j(X)+X^{2^{j-1}-1}N(X),\end{aligned} \tag{5.13}$$

where

$$N(X)=\int_0^1|H(\alpha;X)|^{2^{j-1}}\sum_{1\leqslant|x|\leqslant X}|h(bx\alpha;X)|^{2^{j-1}}d\alpha. \tag{5.14}$$

By orthogonality, it follows from (5.14) that $N(X)$ is equal to the number of integral solutions of the equation

$$bx\sum_{i=1}^{2^{j-2}}(y_i^{k-1}-z_i^{k-1})=\sum_{i=1}^{2^{j-2}}(\Phi(u_i,v_i)-\Phi(t_i,w_i)), \tag{5.15}$$

with $1\leqslant |x|\leqslant X$, and with each of y_i,z_i,u_i,v_i,t_i,w_i $(1\leqslant i\leqslant 2^{j-2})$ bounded in absolute value by X. Let $N_0(X)$ denote the number of such solutions of (5.15) with the right hand side of the equation equal to zero, and let $N_1(X)$ denote the corresponding number of solutions with the latter expression non-zero. Then one has

$$N(X)=N_0(X)+N_1(X). \tag{5.16}$$

We first estimate $N_0(X)$. On considering the underlying diophantine equations, and recalling (5.4), we have

$$N_0(X)\ll XI_j(X)\int_0^1 |h(\alpha;X)|^{2^{j-1}}\,d\alpha.$$

Then on recalling the classical version of Hua's Lemma (see Vaughan [27, Lemma 2.5]), we find that when $2\leqslant j\leqslant k-1$ one has

$$N_0(X)\ll X^{2^{j-1}-j+2+\varepsilon}I_j(X). \tag{5.17}$$

In order to estimate $N_1(X)$ we require some notation. For each integer h, we denote by $r(n;h)$ the number of representations of the integer n in the form

$$n=h\sum_{i=1}^{2^{j-2}}(y_i^{k-1}-z_i^{k-1}),$$

with $|y_i|\leqslant X$ and $|z_i|\leqslant X$ $(1\leqslant i\leqslant 2^{j-2})$. Similarly, for each integer n we take $R(n)$ to be the number of representations of n in the form

$$n=\sum_{i=1}^{2^{j-2}}(\Phi(u_i,v_i)-\Phi(t_i,w_i)),$$

with each of u_i,v_i,t_i,w_i $(1\leqslant i\leqslant 2^{j-2})$ bounded in absolute value by X. Then it follows that

$$N_1(X)\leqslant \sum_{1\leqslant |n|\leqslant 2^{j-1}X^k} R(bn)\sum_{\substack{h|n\\ |h|\leqslant X}} r(n;h).$$

Consequently, on recalling an elementary estimate for the divisor function, we deduce from Cauchy's inequality that

$$\begin{aligned}N_1(X)&\leqslant\Big(\sum_{n\in\mathbb{Z}}R(bn)^2\Big)^{1/2}\Big(\sum_{1\leqslant |n|\leqslant 2^{j-1}X^k}\Big(\sum_{\substack{h|n\\ |h|\leqslant X}} r(n;h)\Big)^2\Big)^{1/2}\\ &\ll X^\varepsilon\Big(\sum_{n\in\mathbb{Z}}R(n)^2\Big)^{1/2}\Big(\sum_{n\in\mathbb{Z}}\sum_{1\leqslant |h|\leqslant X} r(n;h)^2\Big)^{1/2}.\end{aligned} \tag{5.18}$$

But on considering the underlying diophantine equations, it follows from (5.18) that

$$N_1(X)\ll X^\varepsilon(I_{j+1}(X))^{1/2}\Big(X\int_0^1 |h(\alpha;X)|^{2^j}\,d\alpha\Big),$$

whence, on recalling the classical version of Hua's Lemma (see Vaughan [27, Lemma 2.5]), we may conclude that when $2\leqslant j\leqslant k-1$ one has

$$N_1(X)\ll X^\varepsilon(I_{j+1}(X))^{1/2}(X^{2^j-j+1+\varepsilon})^{1/2}. \tag{5.19}$$

On combining (5.13), (5.16), (5.17) and (5.19), we arrive at the estimate

$$I_{j+1}(X)\ll (X^{2^{j-1}}+X^{2^j-j+1+\varepsilon})I_j(X)+X^{2^j-\frac{1}{2}(j+1)+\varepsilon}(I_{j+1}(X))^{1/2}.$$

Consequently, for $2\leqslant j\leqslant k-1$ it follows from our inductive hypothesis (5.12) that

$$I_{j+1}(X)\ll X^{2^{j+1}-j-1+\varepsilon}+X^{2^j-\frac{1}{2}(j+1)+\varepsilon}(I_{j+1}(X))^{1/2},$$

whence the estimate (5.12) follows with j replaced by $j+1$. The conclusion of the lemma therefore follows by induction.

6. Hua's Lemma for binary forms: the inductive step

We are prepared now for our main assault on the proof of Theorem 2. We consider a non-degenerate binary form $\Phi(x,y)$ of the shape (5.1), and we define the exponential sum $H(\alpha;X)$ as in (5.2). In this section we aim to bound $I_j(X)$ for $1\leqslant j\leqslant k-1$, and so in view of Lemma 5.1 we may suppose without loss of generality that $k\geqslant 4$. In view of Lemmata 5.2 and 5.3, moreover, together with the discussion preceding the latter lemma, it suffices to establish Theorem 2 only in the cases where $\Phi(x,y)$ takes the shape (5.1) with $2\leqslant t\leqslant k-2$. We henceforth suppose that the latter is indeed the case.

The following lemma establishes the second estimate recorded in Theorem 2.

Lemma 6.1. *Let $\Phi(x,y)\in\mathbb{Z}[x,y]$ be a non-degenerate form of degree $k\geqslant 5$ of the shape discussed above, and let X be a large real number. Then for each integer j with $1\leqslant j\leqslant k-1$, and for each positive number ε, one has*

$$I_j(X)\ll X^{2^j-j+\frac{1}{2}+\varepsilon}.$$

Proof. The conclusion of Lemma 6.1 when $j=1,2$ is established in Lemma 5.1 above. Hence we restrict attention to those j with $3\leqslant j\leqslant k-1$, and aim to show that for each such j one has

$$I_j(X)\ll X^{2^{j-1}-1}I_{j-1}(X)+X^{2^j-j+\frac{1}{2}+\varepsilon}. \tag{6.1}$$

The conclusion of the lemma plainly follows from (6.1) by induction, with the estimate $I_2(X)\ll X^{2+\varepsilon}$ providing the basis for the induction.

Suppose that j is an integer with $3\leqslant j\leqslant k-1$. We begin our differencing process by noting that, as a consequence of Cauchy's inequality, one has

$$|H(\alpha;X)|^2\ll X\sum_{|u|\leqslant X}|J(\alpha;u;X)|^2, \tag{6.2}$$

where we write

$$J(\alpha;u;X)=\sum_{|v|\leqslant X}e\Big(\alpha\sum_{i=t}^k C_iu^{k-i}v^i\Big), \tag{6.3}$$

and here we have taken the liberty of defining C_t to be B. We divide our treatment into cases according to the value of t.

(a) *Suppose that $t\geqslant j-1$.* It follows from (6.3) by Weyl differencing (see Lemma 3.2) that

$$|J(\alpha;u;X)|^{2^{j-2}}\ll X^{2^{j-2}-j+1}\sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{j-2}|\leqslant 2X}\sum_{x\in\mathfrak{I}_{j-2}}e(\alpha p(x;u;\mathbf{h})), \tag{6.4}$$

where $\mathfrak{I}_{j-2}$ is an interval of integers contained in $[-X,X]$, and the polynomial $p(x;u;\mathbf{h})$ is defined by

$$p(x;u;\mathbf{h})=\sum_{i=t}^k C_iu^{k-i}\Delta_{j-2}(x^i;h_1,\dots,h_{j-2}). \tag{6.5}$$

On applying Hölder's inequality to (6.2), we obtain

$$|H(\alpha;X)|^{2^{j-2}}\ll X^{2^{j-2}-1}\sum_{|u|\leqslant X}|J(\alpha;u;X)|^{2^{j-2}},$$

and hence we deduce from (6.4) that

$$|H(\alpha;X)|^{2^{j-2}}\ll X^{2^{j-1}-j}\mathcal{F}(\alpha), \tag{6.6}$$

where

$$\mathcal{F}(\alpha)=\sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{j-2}|\leqslant 2X}\sum_{|u|\leqslant X}\sum_{x\in\mathfrak{I}_{j-2}}e(\alpha p(x;u;\mathbf{h})). \tag{6.7}$$

On considering the underlying diophantine equations, it follows from (6.5)-(6.7) that

$$\begin{aligned} I_j(X) &\ll X^{2^{j-1}-j}\int_0^1 \mathcal{F}(\alpha)|H(\alpha;X)|^{2^{j-2}}\,d\alpha \\ &\ll X^{2^{j-1}-j}\mathcal{J}(X), \end{aligned} \tag{6.8}$$

where $\mathcal{J}(X)$ denotes the number of integral solutions of the equation

$$p(x;u;\mathbf{h})=\sum_{l=1}^{2^{j-3}}(\Phi(u_l,v_l)-\Phi(t_l,w_l)), \tag{6.9}$$

with $|u|\leqslant X$, $|x|\leqslant X$, and with u_l,v_l,t_l,w_l $(1\leqslant l\leqslant 2^{j-3})$ each bounded in absolute value by X, and $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant j-2)$. We divide the solutions counted by $\mathcal{J}(X)$ into two types. Let $\mathcal{J}_0(X)$ denote the number of solutions of the equation (6.9) counted by $\mathcal{J}(X)$ in which the right hand side of (6.9) is zero, and let $\mathcal{J}_1(X)$ denote the corresponding number of solutions in which the right hand side of (6.9) is non-zero. Then

$$\mathcal{J}(X)=\mathcal{J}_0(X)+\mathcal{J}_1(X). \tag{6.10}$$

We first bound $\mathcal{J}_0(X)$. By hypothesis, one has that C_t is non-zero and that $t\geqslant j-1$, and thus it follows from (6.5) that the polynomial $p(x;u;\mathbf{h})$ is not identically zero. A simple counting argument therefore shows that the number of integral zeros of $p(x;u;\mathbf{h})$, with $|u|\leqslant X$, $|x|\leqslant X$ and $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant j-2)$, is at most $O(X^{j-1})$. Consequently, on considering the underlying diophantine equation, one finds from (6.9) that

$$\mathcal{J}_0(X)\ll X^{j-1}I_{j-1}(X). \tag{6.11}$$

Next we bound $\mathcal{J}_1(X)$. We begin by observing that our hypotheses concerning C_t and j ensure that one may write $p(x;u;\mathbf{h})$ in the form

$$p(x;u;\mathbf{h})=h_1\dots h_{j-2}\sum_{i=t}^{k}D_i\phi_{k-i}(u;\mathbf{h})\psi_i(x;\mathbf{h}), \tag{6.12}$$

where D_i is an integer for $t\leqslant i\leqslant k$, and $D_t\neq 0$, and in which each $\psi_i(x;\mathbf{h})$ is a polynomial with integral coefficients of degree $i-j+2$ with respect to x, and each $\phi_{k-i}(u;\mathbf{h})$ is a polynomial with integral coefficients of degree $k-i$ with respect to u. Indeed, one has $\phi_{k-i}(u;\mathbf{h})=u^{k-i}$ $(t\leqslant i\leqslant k)$, but it is convenient to us for later use to frame $p(x;u;\mathbf{h})$ in a more general shape. In particular, the polynomial $p(x;u;\mathbf{h})$ has degree $k-t\geqslant 2$ with respect to u, and has degree at least one with respect to x.

When n is an integer, denote by $R(n)$ the number of representations of n in the form

$$n=\sum_{l=1}^{2^{j-3}}(\Phi(u_l,v_l)-\Phi(t_l,w_l)),$$

with u_l,v_l,t_l,w_l $(1\leqslant l\leqslant 2^{j-3})$ each bounded in absolute value by X. Similarly, when $h_1,\dots,h_{j-2}$ are integers, denote by $r(n;\mathbf{h})$ the number of representations of n in the form $n=p(x;u;\mathbf{h})$, with $|x|\leqslant X$ and $|u|\leqslant X$. Then in view of (6.9) and (6.12), we find that

$$\mathcal{J}_1(X)=\sum_{n\in\mathbb{Z}\setminus\{0\}}R(n)\sum_{\substack{h_1|n\\ |h_1|\leqslant 2X}}\cdots\sum_{\substack{h_{j-2}|n\\ |h_{j-2}|\leqslant 2X}}r(n;\mathbf{h}).$$

By Cauchy's inequality combined with an elementary estimate for the divisor function, therefore, we obtain

$$\mathcal{J}_1(X)\ll\Bigl(\sum_{n\in\mathbb{Z}}R(n)^2\Bigr)^{1/2}\Bigl(X^\varepsilon\sum_{n\in\mathbb{Z}}\sum_{0<|h_1|\leqslant 2X}\cdots\sum_{0<|h_{j-2}|\leqslant 2X}r(n;\mathbf{h})^2\Bigr)^{1/2}. \tag{6.13}$$

On considering the underlying diophantine equations, it follows from (6.13) that

$$\mathcal{J}_1(X)\ll X^\varepsilon(I_j(X))^{1/2}(\mathcal{K}(X))^{1/2}, \tag{6.14}$$

where $\mathcal{K}(X)$ denotes the number of integral solutions of the equation

$$F(x_1, y_1; \mathbf{h}) = F(x_2, y_2; \mathbf{h}), \tag{6.15}$$

with $|x_i| \leqslant X$, $|y_i| \leqslant X$ $(i = 1, 2)$, and $0 < |h_m| \leqslant 2X$ $(1 \leqslant m \leqslant j - 2)$, where here we write

$$F(x, y; \mathbf{h}) = \sum_{i=t}^{k} D_i \phi_{k-i}(x; \mathbf{h}) \psi_i(y; \mathbf{h}). \tag{6.16}$$

Consider any fixed one of the $O(X^{j-2})$ possible choices for $\mathbf{h}$ satisfying $h_m \neq 0$ $(1 \leqslant m \leqslant j - 2)$. We claim that the polynomial $F(x, y; \mathbf{h})$ is non-degenerate in terms of x and y. In order to verify this claim, we begin by observing that in view of our earlier discussion, one has that the polynomial $F(x, y; \mathbf{h})$ has total degree $k - j + 2$ with respect to x and y, and yet has degree at most $k - t \leqslant k - j + 1$ with respect to x. Thus the partial derivative

$$\frac{\partial^{k-j+2}}{\partial x^{k-j+2}} F(x, y; \mathbf{h}) \tag{6.17}$$

is identically zero. But if $F(x, y; \mathbf{h})$ were degenerate, then for some polynomial $g(t) \in \mathbb{Q}[t]$ of degree $k - j + 2$, and for some integers a, b, c, one would have $F(x, y; \mathbf{h}) = g(ax + by + c)$. Further, since $F(x, y; \mathbf{h})$ is explicit in both x and y, one necessarily has that neither a nor b is zero. Consequently, if $F(x, y; \mathbf{h})$ is degenerate in the manner indicated, then one has

$$\frac{\partial^{k-j+2}}{\partial x^{k-j+2}} F(x, y; \mathbf{h}) = \frac{\partial^{k-j+2}}{\partial x^{k-j+2}} g(ax + by + c) = (k - j + 2)! a^{k-j+2} A_0, \tag{6.18}$$

where A_0 is the leading coefficient of $g(t)$. But the expression (6.18) is non-zero, and this contradicts our earlier observation that the expression (6.17) is identically zero. Thus $F(x, y; \mathbf{h})$ is non-degenerate, as claimed.

Next consider a fixed $\mathbf{h}$ with $h_m \neq 0$ $(1 \leqslant m \leqslant j-2)$. For each integer n we define $w(n)$ to be the number of solutions of the equation $F(x, y; \mathbf{h}) = n$, with $|x| \leqslant X$ and $|y| \leqslant X$. Since $F(x, y; \mathbf{h})$ is non-degenerate, it follows that $w(n) \ll X$ for each n, and moreover one also has

$$\sum_{n \in \mathbb{Z}} w(n) \ll X^2. \tag{6.19}$$

Let $\mathcal{N}$ denote the set of integers n represented in the form $F(x, y; \mathbf{h}) = n$. Then in the notation of the statement of Lemma 4.2, we find that for each fixed $\mathbf{h}$ with $h_m \neq 0$ $(1 \leqslant m \leqslant j - 2)$, the number of solutions of the equation (6.15) with $|x_i| \leqslant X$, $|y_i| \leqslant X$ $(i = 1, 2)$, is bounded above by

$$E(X; \mathcal{N}) \ll X^{1/2+\varepsilon} \sum_{n \in \mathbb{Z}} w(n) + X \max_{n \in \mathbb{Z}} w(n).$$

In view of (6.19), therefore, we conclude that

$$E(X; \mathcal{N}) \ll X^{5/2+\varepsilon},$$

whence

$$\mathcal{K}(X) \ll X^{j-2} \max_{\mathbf{h}} E(X; \mathcal{N}) \ll X^{j+\frac{1}{2}+\varepsilon}. \tag{6.20}$$

On combining (6.8), (6.10), (6.11), (6.14) and (6.20), we arrive at the estimate

$$I_j(X) \ll X^{2^{j-1}-j} \left(X^{j-1} I_{j-1}(X) + \left(X^{j+\frac{1}{2}+\varepsilon} \right)^{1/2} (I_j(X))^{1/2} \right),$$

whence

$$I_j(X) \ll X^{2^{j-1}-1} I_{j-1}(X) + X^{2^j - j + \frac{1}{2}+\varepsilon}.$$

Thus the estimate (6.1) does indeed hold, and so in this case we have established the inductive step.

(b) *Suppose that* $t < j-1$. In this case we are able to follow the trail laid down in the previous case, although now the Weyl differencing is performed in two phases. First, by Weyl differencing $J(\alpha;u;X)$ we deduce from Lemma 3.2 that

$$|J(\alpha;u;X)|^{2^{t-1}} \ll X^{2^{t-1}-t} \sum_{|h_1|\leqslant 2X} \cdots \sum_{|h_{t-1}|\leqslant 2X} \sum_{x\in \mathfrak{I}_{t-1}} e(\alpha \tilde{p}(x;u;\mathbf{h})), \tag{6.21}$$

where $\mathfrak{I}_{t-1}$ is an interval of integers contained in $[-X,X]$, and the polynomial $\tilde{p}(x;u;\mathbf{h})$ is defined by

$$\tilde{p}(x;u;\mathbf{h}) = \sum_{i=t}^{k} C_i u^{k-i} \Delta_{t-1}(x^i;h_1,\dots,h_{t-1}). \tag{6.22}$$

On applying Hölder's inequality to (6.2), we obtain

$$|H(\alpha;X)|^{2^{t-1}} \ll X^{2^{t-1}-1} \sum_{|u|\leqslant X} |J(\alpha;u;X)|^{2^{t-1}},$$

and hence we deduce from (6.21) that

$$|H(\alpha;X)|^{2^{t-1}} \ll X^{2^{t}-t-1} \sum_{|h_1|\leqslant 2X} \cdots \sum_{|h_{t-1}|\leqslant 2X} \sum_{x\in \mathfrak{I}_{t-1}} |K(\alpha;X;\mathbf{h},x)|, \tag{6.23}$$

where

$$K(\alpha;X;\mathbf{h},x) = \sum_{|u|\leqslant X} e(\alpha \tilde{p}(x;u;\mathbf{h})).$$

Next Weyl differencing with respect to the second underlying variable, we deduce from Lemma 3.2 and equation (6.22) that

$$\begin{aligned} &|K(\alpha;X;\mathbf{h},x)|^{2^{j-t-1}} \\ &\qquad \ll X^{2^{j-t-1}-j+t} \sum_{|g_1|\leqslant 2X} \cdots \sum_{|g_{j-t-1}|\leqslant 2X} \sum_{u\in \mathfrak{I}_{j-t-1}} e(\alpha p(x;u;\mathbf{g},\mathbf{h})), \end{aligned} \tag{6.24}$$

where $\mathfrak{I}_{j-t-1}$ is an interval of integers contained in $[-X,X]$, and the polynomial $p(x;u;\mathbf{g},\mathbf{h})$ is now defined by

$$p(x;u;\mathbf{g},\mathbf{h}) = \sum_{i=t}^{k} C_i \Delta_{j-t-1}(u^{k-i};\mathbf{g}) \Delta_{t-1}(x^i;\mathbf{h}). \tag{6.25}$$

Applying Hölder's inequality to (6.23), therefore, we deduce from (6.24) and (6.25) that

$$\begin{aligned} |H(\alpha;X)|^{2^{j-2}} &\ll X^{2^{j-1}-2^{j-t-1}-t} \sum_{|h_1|\leqslant 2X} \cdots \sum_{|h_{t-1}|\leqslant 2X} \sum_{x\in \mathfrak{I}_{t-1}} |K(\alpha;X;\mathbf{h},x)|^{2^{j-t-1}} \\ &\ll X^{2^{j-1}-j} \mathcal{G}(\alpha), \end{aligned} \tag{6.26}$$

where

$$\mathcal{G}(\alpha) = \sum_{|h_1|\leqslant 2X} \cdots \sum_{|h_{j-2}|\leqslant 2X} \sum_{x\in \mathfrak{I}_{t-1}} \sum_{u\in \mathfrak{I}_{j-t-1}} e(\alpha p(x;u;\mathbf{g},\mathbf{h})), \tag{6.27}$$

and here we write

$$\mathbf{h} = (h_1,\dots,h_{t-1}) \quad \text{and} \quad \mathbf{g} = (h_t,\dots,h_{j-2}). \tag{6.28}$$

On considering the underlying diophantine equations, it follows from (6.25)-(6.27) that

$$I_j(X) \ll X^{2^{j-1}-j} \mathcal{J}(X), \tag{6.29}$$

where now $\mathcal{J}(X)$ denotes the number of solutions of the equation

$$p(x;u;\mathbf{g},\mathbf{h}) = \sum_{l=1}^{2^{j-3}} (\Phi(u_l,v_l) - \Phi(t_l,w_l)), \tag{6.30}$$

with $|x|\leqslant X$, $|u|\leqslant X$, and each of u_l, v_l, t_l, w_l $(1\leqslant l\leqslant 2^{j-3})$ bounded in absolute value by X, and with $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant j-2)$, under the same convention regarding $\mathbf{g}$ and $\mathbf{h}$ as in (6.28).

Observe that our hypotheses concerning C_t, j and t ensure that one may write $p(x;u;\mathbf{g},\mathbf{h})$ in the form

$$p(x;u;\mathbf{g},\mathbf{h}) = h_1\dots h_{j-2}\sum_{i=t}^{k} D_i\phi_{k-i}(u;\mathbf{h})\psi_i(x;\mathbf{h}), \tag{6.31}$$

where D_i is an integer for $t\leqslant i\leqslant k$, and $D_t\neq 0$, and in which each $\psi_i(x;\mathbf{h})$ is a polynomial with integral coefficients of degree $i-t+1$ with respect to x, and each $\phi_{k-i}(u;\mathbf{h})$ is a polynomial with integral coefficients of degree $\max\{0,k-i-j+t+1\}$ with respect to u. In particular, the polynomial $p(x;u;\mathbf{g},\mathbf{h})$ has degree $k-j+1\geqslant 2$ with respect to u, and has degree at least one with respect to x. A comparison between (6.31) and (6.12), and the associated discussions, reveals that the polynomials $p(x;u;\mathbf{h})$ and $p(x;u;\mathbf{g},\mathbf{h})$ differ in a manner superficial so far as our method for estimating $\mathcal{J}(X)$ is concerned. Thus, by the argument of part (a) above, mutatis mutandis, one obtains the estimate (6.1) once again, and this completes the proof of the inductive step.

This completes the proof of the lemma.

7. Hua's Lemma for binary forms: the final stages

In order to complete our proof of Theorem 2, we have now only to consider large moments of $H(\alpha;X)$ and establish the final two estimates of the statement of Theorem 2, together with part of the first estimate. This we achieve in two stages, first exploiting Lemma 4.3 with a differencing process of order $k-3$, and finally making use of Theorem 1 via the Hardy-Littlewood method. In order to describe these stages in detail, we require some further notation. When X is a large real number and s is a natural number, define

$$\mathcal{I}_s(X) = \int_0^1 |H(\alpha;X)|^{2s}d\alpha.$$

Lemma 7.1. *Let $\Phi(x,y)\in\mathbb{Z}[x,y]$ be a non-degenerate form of degree $k\geqslant 4$ of the shape discussed in the preamble to the statement of Lemma* 6.1*. Then for each natural number s, and for each positive number ε, one has*

$$\mathcal{I}_{s+2^{k-4}}(X) \ll X^{2^{k-2}-1}\mathcal{I}_s(X) + X^{2^{k-2}-\frac12(k-1)+\varepsilon}(\mathcal{I}_{2s}(X))^{1/2}.$$

Proof. We begin by Weyl differencing $k-3$ times following the pattern of the proof of Lemma 6.1. We divide into cases.

(a) *Suppose that $t=k-2$.* On recalling (6.6) and (6.7), we find that

$$|H(\alpha;X)|^{2^{k-3}} \ll X^{2^{k-2}-k+1}\mathcal{F}(\alpha), \tag{7.1}$$

where

$$\mathcal{F}(\alpha) = \sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{k-3}|\leqslant 2X}\sum_{|u|\leqslant X}\sum_{x\in\mathfrak{I}_{k-3}} e(\alpha p(x;u;\mathbf{h})), \tag{7.2}$$

$$p(x;u;\mathbf{h}) = \sum_{i=k-2}^{k} C_i u^{k-i}\Delta_{k-3}(x^i;\mathbf{h}), \tag{7.3}$$

and $\mathfrak{I}_{k-3}$ is an interval of integers contained in $[-X,X]$. Here we again write C_t in place of B. On considering the underlying diophantine equations, it follows from (7.1)-(7.3) that

$$\begin{aligned}\mathcal{I}_{s+2^{k-4}}(X) &\ll X^{2^{k-2}-k+1}\int_0^1 \mathcal{F}(\alpha)|H(\alpha;X)|^{2s}d\alpha\\ &\ll X^{2^{k-2}-k+1}\mathcal{J}(X),\end{aligned} \tag{7.4}$$

where $\mathcal{J}(X)$ denotes the number of integral solutions of the equation

$$p(x;u;\mathbf{h}) = \sum_{l=1}^{s}(\Phi(u_l,v_l)-\Phi(t_l,w_l)), \tag{7.5}$$

with $|u|\leqslant X$, $|x|\leqslant X$, and with u_l, v_l, t_l, w_l $(1\leqslant l\leqslant s)$ each bounded in absolute value by X, and $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant k-3)$. We divide the solutions counted by $\mathcal{J}(X)$ into two types. Let $\mathcal{J}_0(X)$ denote the number of solutions of the equation (7.5) counted by $\mathcal{J}(X)$ in which the right hand side of (7.5) is zero, and let $\mathcal{J}_1(X)$ denote the corresponding number of solutions in which the right hand side of (7.5) is non-zero. Then

$$\mathcal{J}(X) = \mathcal{J}_0(X) + \mathcal{J}_1(X). \tag{7.6}$$

We first bound $\mathcal{J}_0(X)$. By hypothesis, one has that C_t is non-zero and that $t = k-2$, and thus it follows from (7.3) that the polynomial $p(x;u;\mathbf{h})$ is not identically zero. A simple counting argument therefore shows that the number of integral zeros of $p(x;u;\mathbf{h})$, with $|u|\leqslant X$, $|x|\leqslant X$ and $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant k-3)$, is at most $O(X^{k-2})$. Consequently, on considering the underlying diophantine equations, one finds from (7.5) that

$$\mathcal{J}_0(X) \ll X^{k-2}\mathcal{I}_s(X). \tag{7.7}$$

Next we bound $\mathcal{J}_1(X)$. We begin by observing that our hypotheses concerning C_t and t ensure that $p(x;u;\mathbf{h})$ takes the shape

$$p(x;u;\mathbf{h}) = h_1\dots h_{k-3}F(u,x;\mathbf{h}), \tag{7.8}$$

where

$$F(u,x;\mathbf{h}) = D_2\chi_2(u;\mathbf{h})\omega_1(x;\mathbf{h}) + D_1\chi_1(u;\mathbf{h})\omega_2(x;\mathbf{h}) + D_0\chi_0(u;\mathbf{h})\omega_3(x;\mathbf{h}), \tag{7.9}$$

and D_i is an integer for $i = 0,1,2$, and $D_2 \neq 0$, and in which each $\chi_i(u;\mathbf{h})$ is a polynomial with integral coefficients of degree i with respect to u and $\mathbf{h}$, and each $\omega_i(x;\mathbf{h})$ is a polynomial with integral coefficients of degree i with respect to x and $\mathbf{h}$.

When n is an integer, denote by $R(n)$ the number of representations of n in the form

$$n = \sum_{l=1}^{s}(\Phi(u_l,v_l) - \Phi(t_l,w_l)),$$

with u_l, v_l, t_l, w_l $(1\leqslant l\leqslant s)$ each bounded in absolute value by X. Similarly, when $h_1,\dots,h_{k-3}$ are integers, denote by $r(n;\mathbf{h})$ the number of representations of n in the form $n = p(x;u;\mathbf{h})$, with $|x|\leqslant X$ and $|u|\leqslant X$. Then in view of (7.5) and (7.8), we find that

$$\mathcal{J}_1(X) = \sum_{n\in\mathbb{Z}\setminus\{0\}} R(n) \sum_{\substack{h_1|n\\ |h_1|\leqslant 2X}} \cdots \sum_{\substack{h_{k-3}|n\\ |h_{k-3}|\leqslant 2X}} r(n;\mathbf{h}).$$

By Cauchy's inequality combined with an elementary estimate for the divisor function, therefore, we obtain

$$\mathcal{J}_1(X) \ll \Big(\sum_{n\in\mathbb{Z}} R(n)^2\Big)^{1/2}\Big(X^{\varepsilon}\sum_{n\in\mathbb{Z}\setminus\{0\}}\sum_{0<|h_1|\leqslant 2X}\cdots\sum_{0<|h_{k-3}|\leqslant 2X} r(n;\mathbf{h})^2\Big)^{1/2}.$$

Thus, by considering the underlying diophantine equations, it follows that

$$\mathcal{J}_1(X) \ll X^{\varepsilon}(\mathcal{I}_{2s}(X))^{1/2}(\mathcal{K}(X))^{1/2}, \tag{7.10}$$

where $\mathcal{K}(X)$ denotes the number of integral solutions of the equation

$$F(x_1,y_1;\mathbf{h}) = F(x_2,y_2;\mathbf{h}), \tag{7.11}$$

with $|x_i|\leqslant X$, $|y_i|\leqslant X$ $(i=1,2)$, and $0<|h_m|\leqslant 2X$ $(1\leqslant m\leqslant k-3)$, and subject to the condition that $F(x_i,y_i;\mathbf{h}) \neq 0$ $(i=1,2)$.

We next transform the equation (7.11) into a form amenable to the application of Lemma 4.3. First we rewrite the polynomial $F(x,y;\mathbf{h})$ in the form

$$F(x,y;\mathbf{h}) = \alpha(y;\mathbf{h})x^2 + \beta(y;\mathbf{h})x + \gamma(y;\mathbf{h}), \tag{7.12}$$

where by (7.9) we have that $\alpha(y;\mathbf{h})$ is a non-trivial linear polynomial in y and $\mathbf{h}$ with integral coefficients, and is explicit in y. For a fixed $\mathbf{h}$ with $0<|h_m|\leqslant 2X$ $(1\leqslant m\leqslant k-3)$, let $\mathcal{K}_1(X;\mathbf{h})$ denote the number of solutions of the equation (7.11) counted by $\mathcal{K}(X)$ in which $\alpha(y_i;\mathbf{h})=0$ for $i=1$ or 2. Define the polynomial $\Delta(y;\mathbf{h})$ by

$$\Delta(y;\mathbf{h})=\beta(y;\mathbf{h})^2-4\alpha(y;\mathbf{h})\gamma(y;\mathbf{h}), \tag{7.13}$$

and let $\mathcal{K}_2(X;\mathbf{h})$ denote the corresponding number of solutions of (7.11) in which $\alpha(y_i;\mathbf{h})\neq 0$ $(i=1,2)$, and one has that $\Delta(y;\mathbf{h})$ is identically zero as a polynomial in y. Let $\mathcal{K}_3(X;\mathbf{h})$ denote the corresponding number of solutions in which $\alpha(y_i;\mathbf{h})\neq 0$ $(i=1,2)$, and $\Delta(y;\mathbf{h})$ is not identically zero as a polynomial in y, and moreover one has

$$\alpha(y_2;\mathbf{h})\Delta(y_1;\mathbf{h})=\alpha(y_1;\mathbf{h})\Delta(y_2;\mathbf{h}). \tag{7.14}$$

Finally, let $\mathcal{K}_4(X;\mathbf{h})$ denote the corresponding number of solutions with $\alpha(y_i;\mathbf{h})\neq 0$ $(i=1,2)$, and for which the equation (7.14) does not hold. Then plainly,

$$\mathcal{K}(X)\leqslant \sum_{0<|h_1|\leqslant 2X}\cdots\sum_{0<|h_{k-3}|\leqslant 2X}\sum_{i=1}^{4}\mathcal{K}_i(X;\mathbf{h}). \tag{7.15}$$

We first bound $\mathcal{K}_1(X;\mathbf{h})$. Suppose first that $\alpha(y_i;\mathbf{h})=0$ for $i=1,2$. Since $\alpha(y;\mathbf{h})$ is a non-trivial linear polynomial in y, it follows that for any fixed $\mathbf{h}$ there is at most one permissible choice for $\mathbf{y}$. Since there are trivially $O(X^2)$ possible choices for $\mathbf{x}$, we find that the contribution to $\mathcal{K}_1(X;\mathbf{h})$ from this first class of solutions is $O(X^2)$. Consider next the remaining solutions for which $\alpha(y_i;\mathbf{h})=0$ for at most one value of i. By relabelling variables, we may suppose that $\alpha(y_1;\mathbf{h})=0$. It again follows that for any fixed $\mathbf{h}$ there is at most one permissible choice for y_1 satisfying $\alpha(y_1;\mathbf{h})=0$. Fix any one of the $O(X)$ available choices for x_1, and consider the equation (7.11). Since $\alpha(y_2;\mathbf{h})$ is non-zero, it follows from (7.12) that $F(x_2,y_2;\mathbf{h})$ is explicit in both x_2 and y_2, and consequently a simple counting argument reveals that the number of possible choices for x_2 and y_2 satisfying (7.11) is $O(X)$. Thus there are $O(X^2)$ solutions of this second type, whence

$$\mathcal{K}_1(X;\mathbf{h})\ll X^2. \tag{7.16}$$

Next consider solutions counted by $\mathcal{K}_2(X;\mathbf{h})$. Since $\alpha(y;\mathbf{h})$ is a non-trivial linear polynomial in y, it follows from (7.13) that if $\Delta(y;\mathbf{h})$ is identically zero as a polynomial in y, then $\beta(y;\mathbf{h})$ is divisible by the polynomial $\alpha(y;\mathbf{h})$. Such is immediate when $\gamma(y;\mathbf{h})$ is non-zero, and when $\gamma(y;\mathbf{h})$ is equal to zero one has $\beta(y;\mathbf{h})=0$, and the desired conclusion again follows. But if $\beta(y;\mathbf{h})$ is divisible by $\alpha(y;\mathbf{h})$, then the vanishing of $\Delta(y;\mathbf{h})$ ensures, by (7.13), that $\gamma(y;\mathbf{h})$ is also divisible by $\alpha(y;\mathbf{h})$. We therefore deduce that for some non-zero integers κ_1,κ_2, and some polynomial in y with integral coefficients, say $\delta(y;\mathbf{h})$, one has

$$\kappa_1F(x,y;\mathbf{h})=\alpha(y;\mathbf{h})(\kappa_2x+\delta(y;\mathbf{h}))^2 \tag{7.17}$$

identically as a polynomial in x and y. Let x_2 and y_2 be any one of the $O(X^2)$ permissible choices counted by $\mathcal{K}_2(X;\mathbf{h})$. By hypothesis, one has $F(x_2,y_2;\mathbf{h})\neq 0$. But then it follows from (7.11) and (7.17) that $\alpha(y_1;\mathbf{h})$ and $\kappa_2x_1+\delta(y_1;\mathbf{h})$ are both divisors of the fixed non-zero integer $\kappa_1F(x_2,y_2;\mathbf{h})$. By elementary estimates for the divisor function, therefore, there are at most $O(X^\varepsilon)$ possible choices for integers d_1 and d_2 with $\alpha(y_1;\mathbf{h})=d_1$ and $\kappa_2x_1+\delta(y_1;\mathbf{h})=d_2$. The first of the latter equations uniquely determines y_1, since $\alpha(y;\mathbf{h})$ is explicit and linear in y, and consequently we obtain x_1 uniquely from the second of these equations. Thus we deduce that

$$\mathcal{K}_2(X;\mathbf{h})\ll X^{2+\varepsilon}. \tag{7.18}$$

Consider next the solutions $\mathbf{x}$, $\mathbf{y}$ counted by $\mathcal{K}_3(X;\mathbf{h})$. On the one hand, if the polynomial equation (7.14) is non-trivial in y_1 and y_2, then a simple counting argument shows that there are $O(X)$ permissible choices for y_1 and y_2 satisfying (7.14). Given any one such choice of $\mathbf{y}$, in view of the presumed non-vanishing of $\alpha(y_i;\mathbf{h})$ $(i=1,2)$, it follows from (7.12) that the equation (7.11) is non-trivial in x_1 and x_2, whence there are $O(X)$ permissible choices of x_1 and x_2 satisfying (7.11). Consequently, the total number of solutions of this type is $O(X^2)$. If, on the other hand, the polynomial equation (7.14) is trivial in y_1 and y_2, then it follows that $\Delta(y;\mathbf{h})$ is a constant multiple of $\alpha(y;\mathbf{h})$. An inspection of (7.13) therefore reveals that $\beta(y;\mathbf{h})$ is divisible by $\alpha(y;\mathbf{h})$. But since $\Delta(y;\mathbf{h})$ is at most linear in y, one finds that $\beta(y;\mathbf{h})$ is identically zero, and that $\gamma(y;\mathbf{h})$ is independent of y. In view of (7.12), the equation (7.11) takes the shape

$$\alpha(y_1;\mathbf{h})x_1^2=\alpha(y_2;\mathbf{h})x_2^2.$$

A comparison between the polynomial $\alpha(y;\mathbf{h})x^2$ and that on the right hand side of (7.17) reveals that we may now apply the argument concluding the treatment of $\mathcal{K}_2(X;\mathbf{h})$ above in order to conclude that the number of solutions of this type is $O(X^{2+\varepsilon})$. Thus we have

$$\mathcal{K}_3(X;\mathbf{h})\ll X^{2+\varepsilon}. \tag{7.19}$$

Finally, we provide a bound for $\mathcal{K}_4(X;\mathbf{h})$. Let $\mathbf{x},\mathbf{y}$ be a solution of (7.11) counted by $\mathcal{K}_4(X;\mathbf{h})$. Then on recalling (7.12), it follows from (7.11) that

$$\begin{aligned}\alpha(y_2;\mathbf{h})\left(2\alpha(y_1;\mathbf{h})x_1+\beta(y_1;\mathbf{h})\right)^2-\alpha(y_1;\mathbf{h})\left(2\alpha(y_2;\mathbf{h})x_2+\beta(y_2;\mathbf{h})\right)^2&\\ =\alpha(y_2;\mathbf{h})\Delta(y_1;\mathbf{h})-\alpha(y_1;\mathbf{h})\Delta(y_2;\mathbf{h}).&\end{aligned} \tag{7.20}$$

By hypothesis, for each of the $O(X^2)$ permissible values of $\mathbf{y}$, one has that the right hand side of (7.20) is a non-zero integer, say N. Fix any one such choice of $\mathbf{y}$, and note that our hypotheses ensure also that $\alpha(y_i;\mathbf{h})\neq 0$ $(i=1,2)$. But by Lemma 4.3, the number of solutions of the equation

$$\alpha(y_2;\mathbf{h})\xi^2-\alpha(y_1;\mathbf{h})\eta^2=N,$$

with ξ and η bounded in absolute value by a fixed power of X, is $O(X^\varepsilon)$. Consequently, the number of possible x_i $(i=1,2)$ is also $O(X^\varepsilon)$, and thus we conclude that

$$\mathcal{K}_4(X;\mathbf{h})\ll X^{2+\varepsilon}. \tag{7.21}$$

On recalling (7.10), (7.15), (7.16), (7.18), (7.19) and (7.21), we find that

$$\mathcal{J}_1(X)\ll(\mathcal{I}_{2s}(X))^{1/2}(X^{k-1+\varepsilon})^{1/2},$$

whence by (7.4), (7.6) and (7.7) we have

$$\mathcal{I}_{s+2^{k-4}}(X)\ll X^{2^{k-2}-1}\mathcal{I}_s(X)+X^{2^{k-2}-\frac{1}{2}(k-1)+\varepsilon}(\mathcal{I}_{2s}(X))^{1/2}.$$

Thus we have established the lemma when $t=k-2$.

(b) *Suppose that* $t<k-2$. On recalling (6.26) and (6.27), we find that

$$|H(\alpha;X)|^{2^{k-3}}\ll X^{2^{k-2}-k+1}\mathcal{G}(\alpha), \tag{7.22}$$

where

$$\mathcal{G}(\alpha)=\sum_{|h_1|\leqslant 2X}\cdots\sum_{|h_{k-3}|\leqslant 2X}\sum_{x\in\mathfrak{I}_{t-1}}\sum_{u\in\mathfrak{I}_{k-2-t}}e(\alpha p(x;u;\mathbf{g},\mathbf{h})), \tag{7.23}$$

$$p(x;u;\mathbf{g},\mathbf{h})=\sum_{i=t}^{k}C_i\Delta_{k-2-t}(u^{k-i};\mathbf{g})\Delta_{t-1}(x^i;\mathbf{h}), \tag{7.24}$$

and $\mathfrak{I}_{t-1}$ and $\mathfrak{I}_{k-2-t}$ are intervals of integers contained in $[-X,X]$, and here we write

$$\mathbf{h}=(h_1,\dots,h_{t-1})\quad\text{and}\quad\mathbf{g}=(h_t,\dots,h_{k-3}). \tag{7.25}$$

On considering the underlying diophantine equations, it follows from (7.22)-(7.24) that

$$\mathcal{I}_{s+2^{k-4}}(X)\ll X^{2^{k-2}-k+1}\mathcal{J}(X), \tag{7.26}$$

where now $\mathcal{J}(X)$ denotes the number of solutions of the equation

$$p(x;u;\mathbf{g},\mathbf{h})=\sum_{l=1}^{s}(\Phi(u_l,v_l)-\Phi(t_l,w_l)),$$

with $|u|\leqslant X$, $|x|\leqslant X$, and each of u_l,v_l,t_l,w_l $(1\leqslant l\leqslant s)$ bounded in absolute value by X, and with $|h_m|\leqslant 2X$ $(1\leqslant m\leqslant k-3)$, under the same convention concerning $\mathbf{g}$ and $\mathbf{h}$ as in (7.25).

In view of the convention (7.25), our hypotheses on C_t and t ensure that the polynomial $p(x;u;\mathbf{g},\mathbf{h})$ has the form

$$p(x;u;\mathbf{g},\mathbf{h})=h_1\dots h_{k-3}F(u,x;\mathbf{h}),$$

where $F(u,x;\mathbf{h})$ has the shape (7.9) discussed in the case (a) above. Thus the argument of the previous section, mutatis mutandis, shows that

$$\mathcal{J}(X)\ll X^{k-2}\mathcal{I}_s(X)+(X^{k-1+\varepsilon})^{1/2}(\mathcal{I}_{2s}(X))^{1/2},$$

and the conclusion of the lemma again follows, by means of (7.26).

Lemma 7.2. *Let $\Phi(x,y)\in\mathbb{Z}[x,y]$ be a non-degenerate form of degree $k\geqslant 4$ of the form described in the preamble to the statement of Lemma* 6.1*. When $k=4$, for each positive number ε one has*

$$\int_0^1 |H(\alpha;X)|^4 d\alpha \ll X^{5+\varepsilon},$$

and when $k\geqslant 5$, for each positive number ε one has

$$\int_0^1 |H(\alpha;X)|^{\frac{5}{16}2^k} d\alpha \ll X^{\frac{5}{8}2^k-k+1+\varepsilon}.$$

Proof. We begin by considering the situation when $k=4$. By Lemma 5.1 one has $\mathcal{I}_1(X)\ll X^{2+\varepsilon}$, whence by Lemma 7.1 it follows that

$$\mathcal{I}_2(X)\ll X^3\mathcal{I}_1(X)+X^{5/2+\varepsilon}(\mathcal{I}_2(X))^{1/2},$$

and so

$$\mathcal{I}_2(X)\ll X^{5+\varepsilon}.$$

Thus the desired conclusion holds when $k=4$. If, on the other hand, one has $k\geqslant 5$, then we apply Schwarz's inequality in combination with Lemma 6.1 to obtain

$$\begin{aligned}\mathcal{I}_{\frac{3}{16}2^{k-1}}(X) &= \int_0^1 |H(\alpha;X)|^{\frac{3}{16}2^k} d\alpha\\ &\leqslant \Big(\int_0^1 |H(\alpha;X)|^{2^{k-3}} d\alpha\Big)^{1/2}\Big(\int_0^1 |H(\alpha;X)|^{2^{k-2}} d\alpha\Big)^{1/2}\\ &\ll \Big(X^{2^{k-2}-k+\frac{5}{2}+\varepsilon}\Big)^{1/2}\Big(X^{2^{k-1}-k+\frac{3}{2}+\varepsilon}\Big)^{1/2}\\ &\ll X^{\frac{3}{8}2^k-k+2+\varepsilon}.\end{aligned}$$

Consequently, an application of Lemma 7.1 now yields

$$\begin{aligned}\mathcal{I}_{\frac{5}{16}2^{k-1}}(X) &\ll X^{2^{k-2}-1}\mathcal{I}_{\frac{3}{16}2^{k-1}}(X)+X^{2^{k-2}-\frac{1}{2}(k-1)+\varepsilon}\Big(\mathcal{I}_{\frac{3}{8}2^{k-1}}(X)\Big)^{1/2}\\ &\ll X^{\frac{5}{8}2^k-k+1+\varepsilon}+X^{2^{k-2}-\frac{1}{2}(k-1)+\varepsilon}\Big(\mathcal{I}_{\frac{3}{8}2^{k-1}}(X)\Big)^{1/2}.\end{aligned}\tag{7.27}$$

One now observes that by means of a trivial estimate for $H(\alpha;X)$, on considering the underlying diophantine equations, one has

$$\begin{aligned}\mathcal{I}_{\frac{3}{8}2^{k-1}}(X) &= \int_0^1 |H(\alpha;X)|^{\frac{3}{8}2^k} d\alpha\\ &\ll X^{2^{k-3}}\int_0^1 |H(\alpha;X)|^{\frac{5}{16}2^k} d\alpha\\ &= X^{2^{k-3}}\mathcal{I}_{\frac{5}{16}2^{k-1}}(X).\end{aligned}$$

Consequently, in view of (7.27) one has

$$\mathcal{I}_{\frac{5}{16}2^{k-1}}(X)\ll X^{\frac{5}{8}2^k-k+1+\varepsilon}+X^{\frac{5}{16}2^k-\frac{1}{2}(k-1)+\varepsilon}\Big(\mathcal{I}_{\frac{5}{16}2^{k-1}}(X)\Big)^{1/2},$$

and the second conclusion of the lemma follows immediately.

In order to complete the proof of Theorem 2, we apply the Hardy-Littlewood method. In preparation for this treatment, we record an estimate of use to us later in this paper.

Lemma 7.3. *Under the hypotheses of Theorem* 1*, whenever* $1\leqslant q\leqslant X$ *and* $|q\alpha-r|\leqslant X^{1-d}$*, one has*

$$\sum_{1\leqslant x\leqslant X}\sum_{1\leqslant y\leqslant X} e(\alpha\Phi(x,y))\ll X^{2+\varepsilon}(q+X^d|q\alpha-r|)^{-2^{2-d}}. \tag{7.28}$$

Proof. The estimate (7.28) follows immediately from Theorem 1 via a standard argument (see, for example, Davenport and Heilbronn [15] or Exercise 2 of Vaughan [27, Chapter 2]).

Lemma 7.4. *Let* $\Phi(x,y)\in\mathbb{Z}[x,y]$ *be a non-degenerate form of degree* $k\geqslant 3$*. Then when* $k=3$ *or* 4*, one has for each* $\varepsilon>0$ *the upper bound*

$$\int_0^1 |H(\alpha;X)|^{2^{k-1}}d\alpha\ll X^{2^k-k+\varepsilon},$$

and when $k\geqslant 5$*, for each* $\varepsilon>0$ *one has*

$$\int_0^1 |H(\alpha;X)|^{\frac{9}{16}2^k}d\alpha\ll X^{\frac{9}{8}2^k-k+\varepsilon}.$$

Proof. When $r\in\mathbb{Z}$ and $q\in\mathbb{N}$, write

$$\mathfrak{M}(q,r)=\{\alpha\in[0,1)\ :\ |q\alpha-r|\leqslant X^{1-k}\}.$$

Take $\mathfrak{M}$ to be the union of the intervals $\mathfrak{M}(q,r)$ with $0\leqslant r\leqslant q\leqslant X$ and $(r,q)=1$. Note that the intervals occurring in the latter union are disjoint. Also, write $\mathfrak{m}=[0,1)\setminus\mathfrak{M}$. Observe that by Theorem 1 one has

$$\sup_{\alpha\in\mathfrak{m}}|H(\alpha;X)|\ll X^{2-2^{2-k}+\varepsilon}.$$

Moreover, by Lemmata 5.1 and 7.2 one has the estimate

$$\int_0^1 |H(\alpha;X)|^{s_0(k)}d\alpha\ll X^{2s_0(k)-k+1+\varepsilon},$$

where $s_0(k)=2^{k-2}$ for $k=3,4$, and $s_0(k)=\frac{5}{16}2^k$ when $k\geqslant 5$. Thus we deduce that

$$\begin{aligned}\int_{\mathfrak{m}}|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha &\ll \Bigl(\sup_{\alpha\in\mathfrak{m}}|H(\alpha;X)|\Bigr)^{2^{k-2}}\int_0^1|H(\alpha;X)|^{s_0(k)}d\alpha\\ &\ll X^{2s_0(k)+2^{k-1}-k+\varepsilon}.\end{aligned} \tag{7.29}$$

Next, by (7.28) one has

$$\begin{aligned}&\int_{\mathfrak{M}}|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha\\ &\quad\ll X^{2s_0(k)+2^{k-1}+\varepsilon}\sum_{1\leqslant q\leqslant X}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{|\beta|\leqslant(qX^{k-1})^{-1}}(q+X^kq|\beta|)^{-1-s_0(k)2^{2-k}}d\beta.\end{aligned}$$

Since $s_0(k)\geqslant 2^{k-2}$ for each k, it follows that

$$\begin{aligned}\int_{\mathfrak{M}}|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha &\ll X^{2s_0(k)+2^{k-1}-k+\varepsilon}\sum_{1\leqslant q\leqslant X}\sum_{\substack{a=1\\(a,q)=1}}^{q}q^{-2}\\ &\ll X^{2s_0(k)+2^{k-1}-k+2\varepsilon}.\end{aligned} \tag{7.30}$$

Thus, on combining the estimates (7.29) and (7.30), we arrive at the conclusion

$$\begin{aligned}&\int_0^1|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha\\ &\qquad=\int_{\mathfrak{M}}|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha+\int_{\mathfrak{m}}|H(\alpha;X)|^{s_0(k)+2^{k-2}}d\alpha\\ &\qquad\ll X^{2s_0(k)+2^{k-1}-k+\varepsilon},\end{aligned}$$

and the lemma follows immediately.

The proof of Theorem 2 is completed by applying Lemma 2.3 in combination with Lemmata 5.1, 5.2, 5.3, 6.1, 7.2 and 7.4.

8. Estimates stemming from Vinogradov's methods

When d is larger than 12 or so, a trivial argument employing estimates for exponential sums in a single variable based on variants of Vinogradov's methods provides bounds superior to those recorded in Theorems 1 and 2. It is unfortunate that no more efficient method of exploiting Vinogradov's ideas seems to be available to estimate exponential sums over binary forms. Since the estimates stemming from Vinogradov's methods are, in some sense, a trivial application of the latter techniques, we will be brief in our discussion. We first require some notation.

When s and k are natural numbers, and P is a positive real number, we define $J_{s,k}(P)$ to be the number of solutions of the system of diophantine equations

$$\sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \quad (1\leqslant j\leqslant k),$$

with $1\leqslant x_i, y_i\leqslant P$ $(1\leqslant i\leqslant s)$. We say that an exponent $\Delta(s,k)$ is *permissible* whenever the exponent has the property that $J_{s,k}(P) \ll P^{\lambda_{s,k}}$, with

$$\lambda_{s,k} = 2s - \tfrac{1}{2}k(k+1) + \Delta(s,k).$$

It follows easily that any permissible exponent $\Delta(s,k)$ is non-negative, and moreover, without loss of generality, that $\Delta(s,k)\leqslant\frac{1}{2}k(k+1)$. The calculation of the strongest permissible exponents $\Delta(s,k)$ presently attainable is a matter of considerable complexity, and so there seems little point in discussing such bounds in detail within this paper. Instead, we refer the reader to Wooley [29], [30] and forthcoming work of Boklan and Wooley [6] for details on the strongest available permissible exponents. In the present context it suffices to indicate the general shape of the available bounds. On one hand we have the classical permissible exponents

$$\Delta(rk,k) = \tfrac{1}{2}k^2(1-1/k)^r, \tag{8.1}$$

valid for $r,k\in\mathbb{N}$ (see, for example, Vaughan [27, Theorem 5.1]). On the other hand, the more recent estimates of Wooley [29] yield permissible exponents tending to zero essentially twice as fast with respect to r as the exponent (8.1). Thus, when k is sufficiently large one has that the exponents $\Delta(rk,k)$ are permissible, where

$$\Delta(rk,k) = k^2\log k\left(1-\frac{2}{k}(1-1/\log k)\right)^r \quad \text{for} \quad 1\leqslant r\leqslant r_1(k), \tag{8.2}$$

and

$$\Delta(rk,k) = 5(\log k)^3\left(1-\frac{3}{2k}(1-1/k)\right)^{r-r_1(k)} \quad \text{for} \quad r > r_1(k), \tag{8.3}$$

where here we write $r_1(k) = [k(\log k - \log\log k)] + 1$.

Estimates of the type described above lead immediately to improvements on Theorem 2 when d is large.

Theorem 8.1. *Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ is a non-degenerate form of degree $k\geqslant 3$. Let s be a natural number, and suppose that $\Delta(s,k)$ is a permissible exponent. Then for each integer m with $1\leqslant m\leqslant k$ one has*

$$\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y))\Big|^{2s} d\alpha \ll P^{4s-k+\frac{1}{m}\Delta(s-m(m-1)/2,k)}.$$

Proof. In view of Lemma 2.3, it suffices to establish the theorem when $\Phi(x,y)$ takes the shape (5.1) with $A\neq 0$. But by Hölder's inequality one has

$$\begin{aligned}\int_0^1\Big|\sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y))\Big|^{2s} d\alpha &\ll P^{2s-1}\sum_{0\leqslant y\leqslant P}\int_0^1\Big|\sum_{0\leqslant x\leqslant P} e(\alpha\Phi(x,y))\Big|^{2s} d\alpha\\ &\ll P^{2s}\max_{0\leqslant y\leqslant P}\int_0^1\Big|\sum_{0\leqslant x\leqslant P} e(\alpha\Phi(x,y))\Big|^{2s} d\alpha.\end{aligned} \tag{8.4}$$

But by a simple variant of Ford [18, Theorem 1], the integral on the right hand side of (8.4) satisfies, for any integer m with $1\leqslant m\leqslant k$, the inequality

$$\int_0^1 \Bigl| \sum_{0\leqslant x\leqslant P} e(\alpha\Phi(x,y))\Bigr|^{2s} d\alpha \ll P^{2s-k+\frac{1}{m}\Delta(s-m(m-1)/2,k)}. \tag{8.5}$$

The only issue with which one must contend in the application of the latter theorem concerns the uniformity with respect to the parameter y. However, experts in the use of such efficient differencing techniques within Vinogradov's mean value theorem will rapidly circumvent such difficulties. The desired conclusion follows on combining (8.4) and (8.5).

We next turn our attention to analogues of Theorem 1 stemming from Vinogradov's methods.

Theorem 8.2. *Suppose that $\Phi(x,y)\in\mathbb{Z}[x,y]$ is a non-degenerate form of degree $k\geqslant 3$.*
(i) Let $\alpha\in\mathbb{R}$, and let λ be a real number with $0<\lambda\leqslant 1$. Suppose that whenever $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ satisfy $(a,q)=1$ and $|q\alpha-a|\leqslant P^{\lambda-k}$, then one has $q>P^{\lambda}$. Then for $s\geqslant\frac{1}{2}k(k-1)$ and any permissible exponent $\Delta(s,k-1)$, one has

$$\Bigl|\sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y))\Bigr| \ll P^{2-\mu_s(k)+\varepsilon},$$

where

$$\mu_s(k)=\frac{\lambda-\Delta(s,k-1)}{2s}.$$

(ii) Let $\alpha\in\mathbb{R}$, let r be an integer with $1\leqslant r\leqslant\frac{1}{2}k$, and write $\lambda=1-r/k$. Suppose that whenever $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ satisfy $(a,q)=1$ and $|q\alpha-a|\leqslant P^{\lambda-k}$, then one has $q>P^{\lambda}$. Then if s and t are positive integers with $s\geqslant\frac{1}{2}k(k-1)$, and the exponents $\Delta(s,k-1)$ and $\Delta(t,k)$ are permissible, one has

$$\Bigl|\sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y))\Bigr| \ll P^{2+\varepsilon}\left(P^{-\nu_s(k)}+P^{-\rho_t(k)}\right),$$

where

$$\nu_s(k)=\frac{r-\Delta(s,k-1)}{2rs}\quad \textit{and}\quad \rho_t(k)=\frac{k-r(1+\Delta(t,k))}{2tk}.$$

Proof. It follows from Lemma 2.2 that there is a polynomial $\Psi(x,y)\in\mathbb{Z}[x,y]$ satisfying condition $(\mathcal{C})$, a positive integer D depending at most on the coefficients of Φ, and a positive real number X with $X\asymp P$, satisfying the condition that for every real number α one has

$$\Bigl|\sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y))\Bigr| \ll X(\log X)^2 \max_{|v|\leqslant X}\sup_{\beta\in\mathbb{R}} |h_v(\alpha/D;\beta;X)|, \tag{8.6}$$

where

$$h_v(\theta;\beta;X)=\sum_{|u|\leqslant X} e(\theta\Psi(u,v)+\beta u).$$

But since the coefficient of u^k in $\Psi(u,v)$ is non-vanishing, and depends at most on the coefficients of Φ, a simple variant of Vinogradov's method (see, for example, Theorem 5.2 of Vaughan [27] and its proof) yields, under the hypotheses of part (i) of the theorem, the estimate

$$|h_v(\alpha/D;\beta;X)| \ll X^{1-\mu_s(k)+\varepsilon},$$

uniformly in v. The desired conclusion therefore follows immediately from (8.6). In order to establish part (ii) of the theorem, we apply instead a simple variant of the argument of the proof of Theorem 2 of Wooley [31], and the desired conclusion follows in like manner.

Corollary. *Let $\alpha \in \mathbb{R}$, and suppose that there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a,q)=1$, $|\alpha - a/q| \leqslant q^{-2}$ and $P^{1-1/\sqrt{k}} \leqslant q \leqslant P^{k-1+1/\sqrt{k}}$. Then*

$$\Big| \sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y)) \Big| \ll P^{2-\rho(k)+\varepsilon},$$

where, when k is large, one has $\rho(k)^{-1} = \frac{3}{2}k^2(\log k + O(\log\log k))$.

Proof. This conclusion is essentially immediate from Theorem 8.2 on making use of the permissible exponents recorded in (8.2) and (8.3).

We remark that forthcoming work of Boklan and Wooley [6] will enable improvements in the above estimates to be established when k is relatively small. We conclude this section by remarking that when k is large, an application of the Hardy-Littlewood method, based on the use of the above corollary together with Theorem 8.1 and the permissible exponents recorded in (8.2) and (8.3), shows that there exists a natural number $s_0(k)$ such that whenever $s \geqslant s_0(k)$ one has

$$\int_0^1 \Big| \sum_{0\leqslant x,y\leqslant P} e(\alpha\Phi(x,y)) \Big|^{2s} d\alpha \ll P^{4s-k+\varepsilon}.$$

Moreover, one may take $s_0(k) = \frac{1}{2}k^2(\log k + \log\log k + O(1))$.

9. The addition of binary forms

This section is devoted to proving conclusions of the same type as Theorem 3 by means of the Hardy-Littlewood method. Equipped at this point with Theorems 1 and 2, our application of the circle method is essentially routine, and thus we will be brief in our exposition. Since it is no harder to establish, we prove a localised version of Theorem 3 which yields more precise information concerning the density of integral solutions of the equation (1.1). The modifications required to establish the stated version of Theorem 3 are trivial, and will be easily accomplished even by an inexperienced reader.

Let k be a natural number with $k \geqslant 3$, and let s be an integer exceeding $s_0(k)$, where the latter integer is defined as in the statement of Theorem 3. We consider binary forms $\Phi_j(x,y) \in \mathbb{Z}[x,y]$ of degree k for $1 \leqslant j \leqslant s$, the discriminant of each one of which we assume to be non-zero. Since the hypotheses of Theorem 3 permit us to assume that the form $\Phi_1(x_1,y_1) + \cdots + \Phi_s(x_s,y_s)$ is indefinite, we find that there exists $(\boldsymbol{\xi},\boldsymbol{\eta}) \in \mathbb{R}^{2s} \setminus \{\mathbf{0}\}$ with

$$\Phi_1(\xi_1,\eta_1) + \cdots + \Phi(\xi_s,\eta_s) = 0. \tag{9.1}$$

Further, the discriminant of each Φ_j $(1\leqslant j\leqslant s)$ is non-zero, and so it follows that $(\boldsymbol{\xi},\boldsymbol{\eta})$ is a non-singular real solution of the equation (9.1). By homogeneity, moreover, there is no loss of generality in supposing that

$$\max_{1\leqslant j\leqslant s} \{|\xi_j|, |\eta_j|\} \leqslant 1.$$

Let τ be a positive number sufficiently small in terms of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, and define the boxes

$$\mathcal{B}_j = \{(\xi,\eta) \in \mathbb{R}^2 : |\xi - \xi_j| \leqslant \tau \text{ and } |\eta - \eta_j| \leqslant \tau\} \quad (1\leqslant j\leqslant s). \tag{9.2}$$

Further, define $\mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_s$. We aim to obtain an asymptotic formula for the number, $\mathcal{N}(P) = \mathcal{N}_s(P\mathcal{B};\boldsymbol{\Phi})$, of integral solutions of the equation (1.1) satisfying $(\mathbf{x},\mathbf{y}) \in P\mathcal{B} \cap \mathbb{Z}^{2s}$ (where here we abuse notation by organising components in the obvious manner).

Define the exponential sums $f_j(\alpha)$ $(1\leqslant j\leqslant s)$ by

$$f_j(\alpha) = \sum_{(x,y)\in P\mathcal{B}_j} e(\alpha\Phi_j(x,y)) \quad (1\leqslant j\leqslant s).$$

Then by orthogonality one has

$$\mathcal{N}(P) = \int_0^1 f_1(\alpha)f_2(\alpha)\dots f_s(\alpha)d\alpha. \tag{9.3}$$

We obtain an asymptotic formula for $\mathcal{N}(P)$ by means of the Hardy-Littlewood method. Let δ be a positive number with $\delta < 2^{-2-k}$, and write $Q = P^{2^{k-1}\delta}$. Let $\mathfrak{M}$ denote the union of the intervals

$$\mathfrak{M}(q,r) = \{\alpha \in [0,1) \ : \ |\alpha - r/q| \leqslant QP^{-k}\}$$

with $0 \leqslant r \leqslant q \leqslant Q$ and $(r,q)=1$, and let $\mathfrak{m} = [0,1) \setminus \mathfrak{M}$. We observe for later use that the former sets $\mathfrak{M}(q,r)$ are disjoint. Moreover, in view of (9.3),

$$\mathcal{N}(P) = \int_{\mathfrak{M}} f_1(\alpha) \dots f_s(\alpha) d\alpha + \int_{\mathfrak{m}} f_1(\alpha) \dots f_s(\alpha) d\alpha. \tag{9.4}$$

The evaluation of the minor arc contribution is straightforward. Suppose that $\alpha \in \mathfrak{m}$. By Dirichlet's theorem there exist $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(r,q)=1$, $1 \leqslant q \leqslant Q^{-1}P^k$, and satisfying $|q\alpha - r| \leqslant QP^{-k}$. Further, since $\alpha \notin \mathfrak{M}$, one necessarily has $q > Q$. Thus it follows directly from Theorem 1 that for $1 \leqslant j \leqslant s$, one has

$$\sup_{\alpha \in \mathfrak{m}} |f_j(\alpha)| \ll P^{2+\varepsilon} Q^{-2^{2-k}} = P^{2-2\delta+\varepsilon}.$$

We therefore deduce from Theorem 2 that for each j with $1 \leqslant j \leqslant s$, one has

$$\begin{aligned} \int_{\mathfrak{m}} |f_j(\alpha)|^s d\alpha &\leqslant \Big(\sup_{\alpha \in \mathfrak{m}} |f_j(\alpha)| \Big)^{s-s_0(k)} \int_0^1 |f_j(\alpha)|^{s_0(k)} d\alpha \\ &\ll P^{2s-k-\delta}, \end{aligned}$$

whence by Hölder's inequality,

$$\begin{aligned} \int_{\mathfrak{m}} f_1(\alpha) \dots f_s(\alpha) d\alpha &\ll \prod_{j=1}^{s} \Big(\int_{\mathfrak{m}} |f_j(\alpha)|^s d\alpha \Big)^{1/s} \\ &\ll P^{2s-k-\delta}. \end{aligned} \tag{9.5}$$

In order to analyse the major arc contribution we introduce some further notation. When $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, we write

$$S_j(q,r) = \sum_{x=1}^{q} \sum_{y=1}^{q} e\left(\frac{r}{q} \Phi_j(x,y) \right) \quad (1 \leqslant j \leqslant s),$$

and when $\beta \in \mathbb{R}$ we write

$$v_j(\beta) = \iint_{P\mathcal{B}_j} e(\beta \Phi_j(\xi,\eta)) d\xi d\eta \quad (1 \leqslant j \leqslant s).$$

A straightforward partial summation argument (see, for example, the proof of Lemma 2.7 of Vaughan [27]) shows that whenever $r \in \mathbb{Z}$ and $q \in \mathbb{N}$, then one has

$$f_j(\alpha) - q^{-2} S_j(q,r) v_j(\alpha - r/q) \ll P(q + P^k |q\alpha - r|). \tag{9.6}$$

When $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $0 \leqslant r \leqslant q \leqslant Q$ and $(r,q)=1$, define $f_j^*(\alpha)$ for $|\alpha - r/q| \leqslant QP^{-k}$ by

$$f_j^*(\alpha) = q^{-2} S_j(q,r) v_j(\alpha - r/q) \quad (1 \leqslant j \leqslant s),$$

and define $f_j^*(\alpha)$ to be zero otherwise. Then it follows from (9.6) that whenever $\alpha \in \mathfrak{M}(q,r) \subseteq \mathfrak{M}$, one has

$$f_j(\alpha) - f_j^*(\alpha) \ll PQ^2. \tag{9.7}$$

Since the measure of $\mathfrak{M}$ is $O(Q^3 P^{-k})$, it follows from (9.7) together with a trivial estimate for the $f_i(\alpha)$ that

$$\int_{\mathfrak{M}} f_1(\alpha) \dots f_s(\alpha) - f_1^*(\alpha) \dots f_s^*(\alpha) d\alpha \ll (Q^5 P^{1-k})(P^2)^{s-1} \ll P^{2s-k-\delta}. \tag{9.8}$$

On combining (9.4), (9.5) and (9.8), we find that

$$\mathcal{N}(P) = \mathcal{C}(Q,P)\mathfrak{S}(Q) + O(P^{2s-k-\delta}), \tag{9.9}$$

where for each positive number R we write

$$\mathcal{C}(R,P) = \int_{-RP^{-k}}^{RP^{-k}} v_1(\beta)\dots v_s(\beta)d\beta \tag{9.10}$$

and

$$\mathfrak{S}(R) = \sum_{1\leqslant q\leqslant R} \sum_{\substack{r=1\\(r,q)=1}}^{q} q^{-2s}S_1(q,r)\dots S_s(q,r). \tag{9.11}$$

We next analyse the singular integral $\mathcal{C}^*$, which we define by

$$\mathcal{C}^* = \int_{-\infty}^{\infty} v_1(\beta)\dots v_s(\beta)d\beta. \tag{9.12}$$

Note first that by the argument of the proof of Lemma 2.7 of Brüdern and Wooley [9], for each j one has

$$v_j(\beta) \ll P^2(1+P^k|\beta|)^{-2/k}.$$

Then we deduce from (9.10) and (9.12) that for each positive number R one has

$$\begin{aligned}\mathcal{C}^* - \mathcal{C}(R,P) &\ll P^{2s}\int_{RP^{-k}}^{\infty}(1+P^k\beta)^{-2s/k}d\beta\\ &\ll P^{2s-k}R^{-1/k}.\end{aligned}$$

It follows in particular that the singular integral $\mathcal{C}^*$ is absolutely convergent, and moreover that

$$|\mathcal{C}^*| \ll P^{2s-k} \quad \text{and} \quad \mathcal{C}^* - \mathcal{C}(Q,P) \ll P^{2s-k-\delta}. \tag{9.13}$$

Next, by making a change of variables one obtains

$$\mathcal{C}^* = P^{2s-k}\mathcal{C}, \tag{9.14}$$

where

$$\mathcal{C} = \int_{-\infty}^{\infty}\int_{\mathcal{B}} e(\beta(\Phi_1(\xi_1,\eta_1)+\dots+\Phi_s(\xi_s,\eta_s)))d\boldsymbol{\xi}d\boldsymbol{\eta}d\beta.$$

In view of our hypothesis that $(\boldsymbol{\xi},\boldsymbol{\eta})$ is a non-singular solution of (9.1), a standard application of Fourier's integral formula (see, for example, Lemma 6.2 of Davenport [11]) shows that $\mathcal{C}>0$, and indeed that $\mathcal{C}$ is equal to the volume of the $(2s-1)$-dimensional hypersurface determined by the equation (1.1) contained in the box $\mathcal{B}$.

We define the singular series $\mathfrak{S}$ by

$$\mathfrak{S} = \sum_{q=1}^{\infty} q^{-2s} \sum_{\substack{r=1\\(r,q)=1}}^{q} S_1(q,r)\dots S_s(q,r). \tag{9.15}$$

We may analyse the singular series cheaply by noting that for each j, whenever $(q,r)=1$ it follows from Theorem 1 that

$$S_j(q,r) \ll q^{2-2^{2-k}+\varepsilon}. \tag{9.16}$$

On substituting (9.16) into (9.15), and recalling that $s>2^{k-1}$, we deduce that

$$\mathfrak{S} \ll \sum_{q=1}^{\infty} q^{1-s2^{2-k}+\varepsilon} \ll \sum_{q=1}^{\infty} q^{-1-2^{1-k}} \ll 1. \tag{9.17}$$

Thus $\mathfrak{S}$ converges absolutely, and moreover, on substituting (9.16) into (9.11) we find that

$$\mathfrak{S}-\mathfrak{S}(Q)\ll\sum_{q>Q}q^{-1-2^{-1-k}}\ll Q^{-2^{1-k}}\ll P^{-\delta}. \tag{9.18}$$

On collecting together (9.9), (9.13), (9.14), (9.17) and (9.18), therefore, we may conclude thus far that

$$\mathcal{N}(P)=\mathcal{C}\mathfrak{S}P^{2s-k}+O(P^{2s-k-\delta}), \tag{9.19}$$

where $\mathcal{C}>0$ is the aforementioned $(2s-1)$-volume.

In order to complete our proof of Theorem 3, we have only to conclude our analysis of $\mathfrak{S}$. As is familiar with complete exponential sums of arithmetic type, the sum $S_j(q,r)$ has the quasi-multiplicative property that whenever $(q_1q_2,r)=(q_1,q_2)=1$, then one has

$$S(q_1q_2,r)=S(q_1,rq_2^{k-1})S(q_2,rq_1^{k-1}).$$

Thus a standard argument (see for example §2.6 of Vaughan [27]) shows that the function $S(q)$, which we define by

$$S(q)=q^{-2s}\sum_{\substack{r=1\\(r,q)=1}}^{q}S_1(q,r)\dots S_s(q,r), \tag{9.20}$$

is a multiplicative function of q. In view of the absolute convergence of the series $\mathfrak{S}$, one therefore finds that

$$\mathfrak{S}=\prod_p v_p, \tag{9.21}$$

where the product is over prime numbers, and

$$v_p=\sum_{h=0}^{\infty}S(p^h). \tag{9.22}$$

But by (9.16), (9.20) and (9.22), one has that for each prime number p,

$$v_p-1\ll\sum_{h=1}^{\infty}(p^h)^{-1-2^{1-k}}\ll p^{-1-2^{1-k}}.$$

Then on taking p_0 to be a sufficiently large constant depending at most on the coefficients of the Φ_j, one has

$$\mathfrak{S}=\Big(\prod_{p\leqslant p_0}v_p\Big)\Big(\prod_{p>p_0}(1+O(p^{-1-2^{1-k}}))\Big)=\Big(\prod_{p\leqslant p_0}v_p\Big)(1+O(p_0^{-2^{1-k}})). \tag{9.23}$$

Next we observe that when p is a prime number and h is a natural number, it follows from a standard argument (see, for example, Lemma 2.12 of Vaughan [27]) that

$$\sum_{l=0}^{h}S(p^l)=p^{h(1-2s)}M(p^h),$$

where $M(p^h)$ denotes the number of solutions of the congruence

$$\Phi_1(x_1,y_1)+\dots+\Phi_s(x_s,y_s)\equiv 0\pmod{p^h},$$

with $1\leqslant x_i,y_i\leqslant p^h$ $(1\leqslant i\leqslant s)$. Thus we find that

$$v_p=\lim_{h\to\infty}p^{h(1-2s)}M(p^h), \tag{9.24}$$

and so Theorem 3 follows in all essentials from (9.19), (9.21) and (9.24). In order to establish the corollary to Theorem 3, we note that by Davenport and Lewis [16], for any integers $a_1,\dots,a_t$, the equation

$$a_1x_1^k+\dots+a_tx_t^k=0$$

possesses a non-trivial p-adic solution provided only that $t > k^2$. If we set $x_i = \lambda_i y_i$ $(1\leqslant i\leqslant s)$ in (1.1), then it follows that for each choice of $\boldsymbol{\lambda}$, the equation (1.1) possesses a non-trivial p-adic solution provided only that $s > k^2$. Since the Φ_j have non-zero discriminants, any such non-trivial p-adic solution provides a non-singular p-adic solution to (1.1), and thus it follows easily that for every prime p,

$$M(p^h) \gg p^{h(1-2s)}.$$

We therefore deduce from (9.24) that $v_p > 0$ for every prime number p, whence from (9.23) one has $\mathfrak{S} \gg 1$. On recalling (9.17), (9.19), and our earlier conclusion that $\mathcal{C} > 0$, therefore, we find that

$$P^{2s-k} \ll \mathcal{N}(P) \ll P^{2s-k},$$

whence the corollary also follows.

10. Distribution modulo one and binary forms

Experts will perceive that a pedestrian application of Theorem 1 will fail to achieve the exponent claimed in Theorem 4. We must therefore engineer the kind of modification of Weyl's inequality described, for example, in Baker [3, §3.5]. We begin by recalling a lemma on reciprocal sums.

Lemma 10.1. *Suppose that δ is a positive number, and that α and β are real numbers. Let N, R and B be positive real numbers with $B \gg N^{1+\delta} + R^{1+\delta}$. Suppose further that*

$$\sum_{1\leqslant z\leqslant R} \min\{N, \|z\alpha+\beta\|^{-1}\} \gg B.$$

Then there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with

$$(a,q) = 1, \quad 1\leqslant q \ll NRB^{-1} \quad \text{and} \quad |q\alpha - a| < N^\delta B^{-1}.$$

Proof. This is Lemma 3.3 of Baker [3].

Next we establish an analogue of Weyl's inequality.

Lemma 10.2. *Let k be an integer with $k\geqslant 3$ and let $\Phi(x,y) \in \mathbb{Z}[x,y]$ be a non-degenerate homogeneous form of degree k. When P is a large real number, define the exponential sum $F(\alpha;P)$ by*

$$F(\alpha;P) = \sum_{1\leqslant x\leqslant P}\sum_{1\leqslant y\leqslant P} e(\alpha\Phi(x,y)). \tag{10.1}$$

Suppose that δ is a positive number, and that L and A are positive real numbers with

$$L\leqslant P^k \quad \text{and} \quad A \gg P^{2^{k-1}-1+\delta}L.$$

Suppose further that

$$\sum_{1\leqslant l\leqslant L} |F(l\alpha;P)|^{2^{k-2}} \gg A. \tag{10.2}$$

Then there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with

$$(a,q) = 1, \quad 1\leqslant q\leqslant LP^{2^{k-1}+\delta}A^{-1} \quad \text{and} \quad |q\alpha - a|\leqslant P^{2^{k-1}-k+\delta}A^{-1}. \tag{10.3}$$

Proof. By Lemma 2.2 there is a condensation Ψ of Φ, a non-zero integer D depending at most on the coefficients of Φ, and a positive real number X with $X \asymp P$, satisfying the property that for every real number α one has

$$|F(\alpha;P)| \ll (\log X)^2 \sup_{\beta,\gamma\in\mathbb{R}} |H(\alpha/D;\beta,\gamma;X)|, \tag{10.4}$$

where $H(\theta;\beta,\gamma;X)$ is given by (2.4). We divide our argument into cases.

Suppose first that $\Psi(u,v)$ takes the shape (2.3) with $2\leqslant t\leqslant k-1$. In this case the differencing argument leading to (3.16) shows that for all real numbers α,β,γ, one has

$$|H(\alpha/D;\beta,\gamma;X)|^{2^{k-2}}\ll X^{2^{k-1}-k+\varepsilon}\Big(X^{k-1}+\sum_{1\leqslant h\leqslant G}\min\{X,\|h\alpha/D\|^{-1}\}\Big),$$

where $G=t!(k-t)!|B|t(2X)^{k-1}$. On recalling the hypothesis (10.2), it therefore follows from (10.4) together with a simple estimate for the divisor function that

$$\begin{aligned} A&\ll \sum_{1\leqslant l\leqslant L}|F(l\alpha;P)|^{2^{k-2}}\\ &\ll X^{2^{k-1}-1+\varepsilon}L+X^{2^{k-1}-k+\varepsilon}\sum_{1\leqslant g\leqslant GL}\min\{X,\|g\alpha/D\|^{-1}\}. \end{aligned}\tag{10.5}$$

By hypothesis, one has

$$A\gg P^{2^{k-1}-1+\delta}L\gg X^{2^{k-1}-1+\delta}L.\tag{10.6}$$

Then since we may suppose that $\delta\geqslant 3\varepsilon$, it follows from (10.5) that

$$\sum_{1\leqslant g\leqslant GL}\min\{X,\|g\alpha/D\|^{-1}\}\gg AX^{k-2^{k-1}-\varepsilon}.$$

But (10.6) yields

$$AX^{k-2^{k-1}-\varepsilon}\gg X^{k-1+\delta-\varepsilon}L\gg X^{1+\eta}+(GL)^{1+\eta},$$

where $\eta=\delta/(2k)$, and thus it follows from Lemma 10.1 that there exist $b\in\mathbb{Z}$ and $r\in\mathbb{N}$ with $(b,r)=1$,

$$1\leqslant r\ll X^kL(AX^{k-2^{k-1}-\varepsilon})^{-1}\quad\text{and}\quad |r\alpha/D-b|<X^{\eta}(AX^{k-2^{k-1}-\varepsilon})^{-1}.$$

Write $q=r/(r,D)$ and $a=bD/(r,D)$. Then we may conclude that there exist $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ with $(a,q)=1$,

$$1\leqslant q\ll LX^{2^{k-1}+\varepsilon}A^{-1}\ll LP^{2^{k-1}+\delta/2}A^{-1}$$

and

$$|q\alpha-a|\ll X^{2^{k-1}-k+\eta+\varepsilon}A^{-1}<P^{2^{k-1}-k+\delta/2}A^{-1}.$$

The first case of the lemma follows immediately.

Suppose next that $\Psi(u,v)$ takes the shape (2.3) with $t=k$, so that for fixed integers A and B depending at most on the coefficients of Φ, one has $\Psi(u,v)=Au^k+Bv^k$. This is essentially the classical (diagonal) situation. But now the argument of §3 leading to (3.6) via (3.2) and (3.4) shows that for all real numbers α,β,γ, one has

$$|H(\alpha/D;\beta,\gamma;X)|^{2^{k-2}}\ll X^{2^{k-1}-k+\varepsilon}\Big(X^{k-1}+\sum_{1\leqslant h\leqslant G}\min\{X,\|h\alpha/D\|^{-1}\}\Big),$$

where now $G=\max\{|A|,|B|\}k!(2X)^{k-1}$. Then on recalling the hypothesis (10.2), it follows from (10.4) together with a simple estimate for the divisor function that (10.5) again holds. It is therefore apparent that the argument of the previous case, mutatis mutandis, again establishes the existence of $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ satisfying (10.3). This completes the proof of the lemma.

We are now equipped to prove Theorem 4. With the hypotheses of the statement of Theorem 4, suppose, if possible, that with some large number N one has

$$\|\alpha\Phi(m,n)\|>N^{\delta-2^{2-d}}$$

for $1\leqslant m,n\leqslant N$. On recalling the notation (10.1), it follows from Harman [19, Lemma 5] that with $L=[N^{2^{2-d}-\delta}]+1$, one has

$$\sum_{1\leqslant l\leqslant L}|F(l\alpha;N)|>\frac{N^2}{6}.$$

Then by Hölder's inequality,

$$\sum_{1\leqslant l\leqslant L}|F(l\alpha;N)|^{2^{d-2}}\gg A,$$

where we write $A=N^{2^{d-1}}L^{1-2^{d-2}}$. On checking the hypotheses of Lemma 10.2, we may conclude that there exist $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ with

$$(a,q)=1,\quad 1\leqslant q\leqslant LN^{2^{d-1}+\delta}A^{-1}\quad\text{and}\quad |q\alpha-a|\leqslant N^{2^{d-1}-d+\delta}A^{-1}.$$

Consequently, one has $1\leqslant q\leqslant N^{1-2\delta/3}$, whence

$$\begin{aligned}\|\alpha\Phi(q,q)\|&\leqslant|\Phi(1,1)|q^{d-1}|q\alpha-a|\\&\leqslant|\Phi(1,1)|L^{2^{d-2}-1}N^{-1-\delta/4}\\&\leqslant L^{-1}\leqslant N^{\delta-2^{2-d}}.\end{aligned}$$

This completes the proof of Theorem 4.

References

1. G. I. Arkhipov and A. A. Karatsuba, *A multidimensional analogue of the Waring problem*, Dokl. Akad. Nauk SSSR **295** (1987), 521–523. (Russian)
2. G. I. Arkhipov, A. A. Karatsuba and V. N. Chubarikov, *Theory of multiple trigonometric sums*, "Nauka", Moscow, 1987.
3. R. C. Baker, *Diophantine inequalities, London Mathematical Society Monographs, New Series, 1*, Oxford University Press, New York, 1986.
4. B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263.
5. B. J. Birch and H. Davenport, *Note on Weyl's inequality*, Acta Arith. **7** (1961/62), 273–277.
6. K. D. Boklan and T. D. Wooley, *On Weyl sums for smaller exponents* (in preparation).
7. E. Bombieri and J. Pila, *The number of integral points on arcs and ovals*, Duke Math. J. **59** (1989), 337–357.
8. E. Bombieri and W. M. Schmidt, *On Thue's equation*, Invent. Math. **88** (1987), 69–81.
9. J. Brüdern and T. D. Wooley, *The addition of binary cubic forms*, Philos. Trans. Roy. Soc. London Ser. A **356** (1998), 701–737.
10. S. Chowla and H. Davenport, *On Weyl's inequality and Waring's problem for cubes*, Acta Arith. **6** (1961/62), 505–521.
11. H. Davenport, *Cubic forms in thirty-two variables*, Philos. Trans. Roy. Soc. London Ser. A **251** (1959), 193–232.
12. H. Davenport, *Cubic forms in* 29 *variables*, Proc. Roy. Soc. Ser. A **266** (1962), 287–298.
13. H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, Ann Arbor Publishers, Ann Arbor, 1962.
14. H. Davenport, *Cubic forms in* 16 *variables*, Proc. Roy. Soc. Ser. A **272** (1963), 285–303.
15. H. Davenport and H. Heilbronn, *On Waring's problem: two cubes and one square*, Proc. London Math. Soc. **43** (1937), 73–104.
16. H. Davenport and D. J. Lewis, *Homogeneous additive equations*, Proc. Roy. Soc. Ser. A **274** (1963), 443–460.
17. T. Estermann, *Einige Sätze über quadratfrei Zahlen*, Math. Ann. **105** (1931), 653–662.
18. K. B. Ford, *New estimates for mean values of Weyl sums*, Internat. Math. Res. Notices (1995), 155–171.
19. G. Harman, *Trigonometric sums over primes I*, Mathematika **28** (1981), 249–254.
20. D. R. Heath-Brown, *Cubic forms in ten variables*, Proc. London Math. Soc. (3) **47** (1983), 225–257.
21. C. Hooley, *On nonary cubic forms*, J. Reine Angew. Math. **386** (1988), 32–98.
22. C. Hooley, *On nonary cubic forms. II*, J. Reine Angew. Math. **415** (1991), 95–165.
23. C. Hooley, *On nonary cubic forms. III*, J. Reine Angew. Math. **456** (1994), 53–63.
24. D. J. Lewis, *Diophantine problems: solved and unsolved*, Number Theory and Applications (Banff, AB, 1988), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. (R. A. Mollin, ed.), vol. 265, Kluwer Academic Publishers, Dordrecht, 1989, pp. 103–121.
25. W. M. Schmidt, *The density of integer points on homogeneous varieties*, Acta Math. **154** (1985), 243–296.
26. W. Tartakovsky, *Über asymptotische Gesetze der allgemeinen Diophantischen Analyse mit vielen Unbekannten*, Bull. Acad. Sci. USSR (1935), 483–524.
27. R. C. Vaughan, *The Hardy–Littlewood Method, second edition*, Cambridge University Press, 1997.
28. R. C. Vaughan and T. D. Wooley, *Further improvements in Waring's problem*, Acta Math. **174** (1995), 147–240.
29. T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399.
30. T. D. Wooley, *Quasi-diagonal behaviour in certain mean value theorems of additive number theory*, J. Amer. Math. Soc. **7** (1994), 221–245.
31. T. D. Wooley, *New estimates for Weyl sums*, Quart. J. Math. Oxford (2) **46** (1995), 119–127.
32. T. D. Wooley, *New estimates for smooth Weyl sums*, J. London Math. Soc. (2) **51** (1995), 1–13.

Department of Mathematics, University of Michigan, East Hall, 525 East University Avenue, Ann Arbor, MI 48109-1109, U.S.A.

E-mail address: wooley@math.lsa.umich.edu