ON THE PAUCITY OF NON-DIAGONAL SOLUTIONS IN CERTAIN DIAGONAL DIOPHANTINE SYSTEMS

C. M. Skinner[1] and T. D. Wooley[2]

Dedicated to Professor Christopher Hooley, F. R. S., on the occasion of his 67th birthday

1. Introduction

The investigation of non-trivial solutions of systems of symmetric diagonal equations has a long and distinguished history (see [4, 5]), and raises interesting questions concerning both the density of rational points on algebraic varieties, and the theory of the Hardy-Littlewood method (see [15]). When the number of variables is not too large compared with the underlying degrees, it is believed that the number of diagonal solutions of such systems exceeds the corresponding number of non-diagonal solutions. Let us illustrate this idea with the example which lies at the heart of our paper. Denote by $S_k(P)$ the number of integral solutions of the system

$$\begin{aligned} u_1^k + u_2^k + u_3^k &= v_1^k + v_2^k + v_3^k, \\ u_1 + u_2 + u_3 &= v_1 + v_2 + v_3, \end{aligned} \tag{1.1}$$

with $1 \le u_i, v_i \le P$ $(1 \le i \le 3)$, and let $T(P)$ denote the corresponding number of *trivial* solutions of (1.1), which is to say, the number of solutions for which the u_i are a permutation of the v_j. Plainly, $T(P) = 6P^3 + O(P^2)$, and so the conjecture alluded to above implies that $S_k(P) - T(P) = o(P^3)$. Greaves [7] has very recently established just such an asymptotic formula by pursuing a treatment related to earlier sieving methods of Greaves [6] and Hooley [8, 9, 10, 11]. By using the large sieve in combination with Weil's resolution of the Riemann Hypothesis for curves over finite fields, Greaves [7] has established for $k \ge 4$ the upper bound

$$S_k(P) - T(P) \ll_{\varepsilon,k} P^{17/6+\varepsilon}.$$

In this paper, through an essentially elementary method which avoids any use of sieves or reference to the Riemann Hypothesis for varieties over finite fields, we establish the estimate for $S_k(P) - T(P)$ contained in the following theorem.

---

[1] Supported by an N.S.F. Graduate Fellowship.
[2] Research supported in part by NSF grant DMS-9303505, an Alfred P. Sloan Research Fellowship, and a Fellowship from the David and Lucile Packard Foundation.

Theorem. *When k is an integer exceeding 3, and P is a positive real number,*

$$S_k(P) - T(P) \ll_{\varepsilon,k} P^{\frac{8}{3}+\frac{1}{k-1}+\varepsilon}. \tag{1.2}$$

Consequently, when $k \geq 5$ one has the asymptotic formula

$$S_k(P) = 6P^3 + O_{\varepsilon,k}(P^{\frac{8}{3}+\frac{1}{k-1}+\varepsilon}).$$

We note that the above theorem provides estimates superior to those of Greaves [7] when $k \geq 8$, and provides non-trivial estimates for $k \geq 5$. In the cases $k = 2$ and $k = 3$ the behaviour of $S_k(P)$ is better understood. Rogovskaya [12], improving on work of Bykovskii, has shown that

$$S_2(P) = \frac{18}{\pi^2} P^3 \log P + O(P^3),$$

and Vaughan and Wooley [15, Theorem 1.2] have established the bounds

$$P^2(\log P)^5 \ll S_3(P) - T(P) \ll P^2(\log P)^5.$$

Moreover, as is noted in the introduction of Greaves [7], in the cases $k = 4$ and $k = 5$ the existence of dilations of known integral solutions leads to the lower bounds

$$S_k(P) - T(P) \gg P \quad (k = 4, 5).$$

Indeed, Bremner [3] has found a method which yields large families of parametric solutions of the system (1.1) in the case $k = 5$.

In order to prove our theorem we generalise the efficient slicing method developed in our work [14] on sums of two kth powers. In essence our argument consists of three steps. First we show that the number of solutions of the system (1.1) counted by $S_k(P) - T(P)$ may be estimated in terms of the number of solutions of an auxiliary equation of the shape

$$xF(x, y, h) = G(z, w, h), \tag{1.3}$$

for suitable polynomials F and G, and with each of the variables lying in a box of size comparable to P. Next, in the second step, we use an appropriate version of Siegel's lemma to establish the existence of linearly independent integral vectors (a_i, b_i, c_i, d_i) $(i = 1, 2)$, with components of size about $P^{1/3}$, for which $a_i h + b_i w = c_i z + d_i x$ $(i = 1, 2)$. Except in circumstances in which certain variables are small, it is then possible to eliminate h and w from the equation (1.3) by summing over all possible choices of the slicing variables a_i, b_i, c_i and d_i. Moreover by homogeneity, and by exploiting a coprimality condition which may be imposed on x and z, one deduces from (1.3) that x must be a divisor of a function of the slicing variables alone. The latter implies that x is determined rather efficiently within our slicing argument. Finally, in the third step, we estimate the number of solutions y, z of

(1.3), for fixed values of x and the slicing variables, by appealing to a result of Bombieri and Pila [1, Theorem 5] on the number of integral points on a plane affine curve. In this way we obtain an estimate for $S_k(P) - T(P)$ which is roughly the product of the cardinality of the set of slicing variables, and the number of points on the affine curve resulting from our slicing argument.

It may be possible to generalise our efficient slicing argument so as to obtain non-trivial estimates, when $1 \le k_1 < k_2 < \cdots < k_t$, for the number of integral solutions of the system of equations

$$x_1^{k_j} + \cdots + x_{t+1}^{k_j} = y_1^{k_j} + \cdots + y_{t+1}^{k_j} \quad (1 \le j \le t),$$

with $1 \le x_i, y_i \le P$ $(1 \le i \le t+1)$. Of the three steps outlined above which comprise our basic method, the first may be handled as in the treatment of Wooley [16], and the second through a suitable extension of Lemma 2.2 of this paper. Only the third step, which depends for its success on the absolute irreducibility of a certain polynomial, causes serious difficulties. Nonetheless, we believe that this paper provides a reasonable framework with which to work towards a solution of such problems.

In §2 below we establish the inhomogeneous version of Siegel's Lemma which forms the foundation of our slicing argument. We also record in that section the absolute irreducibility criteria required in our application of the estimate of Bombieri and Pila. Then, in §§3, 4 and 5, we conduct the proof of our theorem according to the plan laid out above, dividing into cases in §§4 and 5 according to whether k is odd or even.

The authors gratefully acknowledge the inspiration they have received from correspondence with George Greaves on the subject of sieve methods and their application to the non-trivial solutions of diagonal equations. We thank also the referee for a number of comments which have improved the exposition of this paper, and in particular, for bringing to our attention a significant oversight in our original proof of Lemma 4.1.

Throughout this paper k will denote a positive integer, and ε will denote a sufficiently small positive number. Also, $\ll$ and $\gg$ denote Vinogradov's well-known notation, and $[x]$ denotes the largest integer not exceeding x.

2. Preliminary lemmata

The first two lemmata of this section are devoted to a proof of an inhomogeneous version of Siegel's Lemma suitable for our application in §§4 and 5. What is needed is a lemma which establishes the existence of two small linearly independent solutions of a linear equation with coefficients which potentially may be of different magnitudes. Thus, while Bombieri and Vaaler [2, Theorem 2] provides such a conclusion of use when the coefficients are of comparable size, we must work a little harder for our slightly sharper conclusion. We start by providing a lower bound for the number of integral solutions of a linear equation inside a normalised box.

Lemma 2.1. *Let $s \geq 2$, and let $Q_1, \dots, Q_s$ be positive real numbers. Suppose that $x_1, \dots, x_s$ are integers with $1 \leq |x_i| \leq Q_i$ $(1 \leq i \leq s)$. Then, for $t \geq 2$, the equation $a_1 x_1 + \dots + a_s x_s = 0$ has at least $[t/2]$ non-zero integral solutions $(a_1, \dots, a_s)$ with*

$$|a_i| \leq Q_i^{-1}(stQ_1 \dots Q_s)^{1/(s-1)} \quad (1 \leq i \leq s). \tag{2.1}$$

Proof. Write $\mathcal{Q} = (stQ_1 \dots Q_s)^{1/(s-1)}$ and $N = \sum_{i=1}^{s} Q_i[\mathcal{Q}/Q_i]$. Let $x_1, \dots, x_s$ be integers satisfying the hypothesis of the lemma. When n is a positive integer, let $r(n; \mathbf{x})$ denote the number of representations of n by the linear expression

$$b_1 x_1 + \dots + b_s x_s, \tag{2.2}$$

with $0 \leq b_i \leq \mathcal{Q}/Q_i$ $(1 \leq i \leq s)$. Since $r(n; \mathbf{x})$ is zero unless $0 \leq |n| \leq N$, we deduce that

$$\sum_{0 \leq |n| \leq N} r(n; \mathbf{x}) = \prod_{i=1}^{s} \left([\mathcal{Q}/Q_i] + 1 \right).$$

But it follows from the equation

$$s^{-1} \sum_{j=1}^{s} \prod_{\substack{i=1 \\ i \neq j}}^{s} \mathcal{Q}/Q_i = t \sum_{j=1}^{s} Q_j,$$

that

$$\sum_{0 \leq |n| \leq N} r(n; \mathbf{x}) > t \sum_{j=1}^{s} Q_j \left([\mathcal{Q}/Q_j] + 1 \right) > \tfrac{1}{2} t(2N + 1).$$

Consequently there exists an integer n with $0 \leq |n| \leq N$ for which $r(n; \mathbf{x}) > \frac{1}{2}t$. Let n_0 be any such integer, and let $\mathbf{b}^{(i)}$ $(1 \leq i \leq [t/2] + 1)$ be distinct s-tuples for which the expression (2.2) represents n_0. When $1 \leq j \leq s$ and $2 \leq i \leq [t/2] + 1$, write $a_j^{(i)} = b_j^{(i)} - b_j^{(1)}$. Then since the $\mathbf{b}^{(i)}$ are distinct, we deduce that the $\mathbf{a}^{(i)}$, for $2 \leq i \leq [t/2]+1$, are distinct non-zero solutions of the equation $a_1 x_1 + \dots + a_s x_s = 0$ satisfying (2.1). This completes the proof of the lemma.

Next we show that inside a normalised box of sufficiently large size, there cannot be too many linearly dependent solutions of a linear equation, whence there exist two “small” linearly independent solutions.

Lemma 2.2. *Suppose that $s \geq 3$, and that $Q_1, \dots, Q_s$ are positive real numbers. Let $x_1, \dots, x_s$ be integers with $1 \leq |x_i| \leq Q_i$ $(1 \leq i \leq s)$. Then there exists a real number X with*

$$1 \leq X \leq (2sQ_1 \dots Q_s)^{1/(s-1)}$$

such that the equations

$$a_1 x_1 + \dots + a_s x_s = b_1 x_1 + \dots + b_s x_s = 0,$$

are soluble in integers a_i, b_i $(1 \le i \le s)$ *with* $\mathbf{a} \ne \pm\mathbf{b}$,

$$(a_1, \dots, a_s) = (b_1, \dots, b_s) = 1,$$

and

$$|a_i| \le Q_i^{-1} X, \quad |b_i| \le Q_i^{-1}(6sQ_1 \dots Q_s/X)^{1/(s-2)} \quad (1 \le i \le s).$$

Proof. Let $\mathcal{Q} = (2sQ_1 \dots Q_s)^{1/(s-1)}$. Also, let $\mathcal{A} = \mathcal{A}(\mathbf{x})$ denote the set of all non-zero integral solutions $\mathbf{a}$ of the equation

$$a_1x_1 + \dots + a_sx_s = 0 \tag{2.3}$$

with $(a_1, \dots, a_s) = 1$. It follows from Lemma 2.1 that such a solution $\mathbf{a}$ exists with $|a_i| \le Q_i^{-1}\mathcal{Q}$ $(1 \le i \le s)$. When $\mathbf{a} \in \mathcal{A}$, define the height function H by

$$H(\mathbf{a}) = \max_{1 \le i \le s} Q_i|a_i|.$$

Let $X = \min_{\mathbf{a}\in\mathcal{A}} H(\mathbf{a})$, and denote by $\mathbf{a}^*$ any fixed element of $\mathcal{A}$ with $H(\mathbf{a}^*) = X$. Also, let $Y = (6sQ_1 \dots Q_s/X)^{1/(s-2)}$. We note that since $1 \le X \le \mathcal{Q}$, we have $Y > \mathcal{Q}$, and hence $H(\mathbf{a}^*) < Y$.

We suppose that for every $\mathbf{b} \in \mathcal{A}$, either $\mathbf{b} = \pm\mathbf{a}^*$, or $H(\mathbf{b}) > Y$. If we derive a contradiction, then the conclusion of the lemma will follow. When Z is a positive number, define $N(Z) = N(Z; \mathbf{x})$ to be the number of non-zero solutions $\mathbf{a}$ of the equation (2.3) with $H(\mathbf{a}) \le Z$, and define $N^*(Z) = N^*(Z; \mathbf{x})$ to be the corresponding number of solutions with $(a_1, \dots, a_s) = 1$. Thus

$$N(Y) = \sum_{1 \le d \le Y} N^*(Y/d). \tag{2.4}$$

Observe that $N^*(Z)$ is zero when $Z < X$, and by hypothesis $N^*(Z) = 2$ when $X \le Z \le Y$ (since, in such circumstances, $N^*(Z)$ counts only the solutions $\mathbf{a} = \pm\mathbf{a}^*$). Then from (2.4),

$$N(Y) = \sum_{1 \le d \le Y/X} 2 \le 2Y/X. \tag{2.5}$$

But an application of Lemma 2.1 with $t = 2(3\mathcal{Q}/X)^{(s-1)/(s-2)}$ yields

$$N(Y) \ge \left[(3\mathcal{Q}/X)^{(s-1)/(s-2)}\right]. \tag{2.6}$$

On combining (2.5) and (2.6) we obtain

$$(3\mathcal{Q}/X)^{(s-1)/(s-2)} - 1 \le 2 \cdot 3^{1/(s-2)}(\mathcal{Q}/X)^{(s-1)/(s-2)},$$

which implies a contradiction. This completes the proof of the lemma.

We now establish estimates for the number of integral points on certain plane affine curves. We require estimates independent of the coefficients of the polynomials defining the curves. Fortunately such estimates are given by Bombieri and Pila [1, Theorem 5], provided that the latter polynomials are absolutely irreducible. In Lemmata 2.3 and 2.5 below we record the necessary criteria for absolute irreducibility. The first criterion stems from work of Schmidt, and the second we establish by a fairly standard argument following closely the proof of [14, Lemma 2.4].

Lemma 2.3. *Let*

$$f(x,y) = g_0 y^d + g_1(x)y^{d-1} + \cdots + g_d(x),$$

be a polynomial with coefficients in a field K, and suppose that g_0 is a non-zero constant. Put

$$\psi(f) = \max_{1\le i\le d} \frac{1}{i}\deg(g_i),$$

and suppose that $\psi(f) = m/d$ with $(m,d) = 1$. Then $f(x,y)$ is absolutely irreducible.

Proof. This is Theorem III.1B of Schmidt [13].

Lemma 2.4. *Let $p(y)$ and $r(x)$ be polynomials with integral coefficients, of respective degrees k and m, and let $q(x,y)$ be a polynomial with integral coefficients of total degree l. Suppose, in addition, that $k > m > l+1$ and $(k,m) = 1$. Then the number, $N(X;p,q,r)$, of solutions of the diophantine equation*

$$p(y) + xyq(x,y) + r(x) = 0,$$

with $0 \le x, y \le X$, satisfies $N(X;p,q,r) \ll_{k,\varepsilon} X^{1/k+\varepsilon}$.

Proof. We apply Lemma 2.3 with $f(x,y) = p(y) + xyq(x,y) + r(x)$, noting that $\psi(f) = \max\{m/k, \alpha\}$, where

$$\alpha \le \max_{0\le j\le l} \frac{l-j+1}{k-j-1} = \frac{l+1}{k-1} < \frac{m}{k}.$$

Thus $\psi(f) = m/k$ with $(m,k) = 1$, and so $p(y) + xyq(x,y) + r(x)$ is absolutely irreducible. It therefore follows from Bombieri and Pila [1, Theorem 5], as in the argument of the proof of [14, Corollary 2.3.1], that

$$N(X;p,q,r) \ll_k X^{1/k} \exp\left(12(k\log X \log\log X)^{1/2}\right),$$

whence the lemma follows immediately.

Lemma 2.5. *When a_1, a_2, b_1 and b_2 are non-zero integers with $a_1 \ne a_2$, define the polynomial $f(u,v,h) = f(u,v,h;\mathbf{a},\mathbf{b})$ by*

$$\begin{aligned} f(u,v,h) = &\left((h-a_1)^k + (h-b_1u)^k - (h-a_1-b_1u)^k\right) \\ &- \left((h-a_2)^k + (h-b_2v)^k - (h-a_2-b_2v)^k\right). \end{aligned}$$

Then when $k \ge 2$, there exists a set $\mathcal{H} = \mathcal{H}(\mathbf{a},\mathbf{b})$, with cardinality at most k^3, with the property that for each fixed non-zero h with $h \notin \mathcal{H}$, the polynomial $f(u,v,h)$ is absolutely irreducible in $\mathbb{Q}[u,v]$.

Proof. We start by observing that without loss of generality, we may suppose that $b_1 = b_2 = 1$. Write $G(\xi_1,\xi_2,\xi_3) = \xi_3^{k-1} f(\xi_1/\xi_3, \xi_2/\xi_3, h)$. Then by using a standard argument, as in the proof of [14, Lemma 2.4], it follows that $f(u,v,h)$ is absolutely irreducible in $\mathbb{Q}[u,v]$ provided that there are no non-trivial solutions $\boldsymbol{\xi} \in \mathbb{C}^3$ of the system

$$G = \frac{\partial G}{\partial \xi_1} = \frac{\partial G}{\partial \xi_2} = 0. \tag{2.7}$$

Suppose that $\boldsymbol{\xi} = \boldsymbol{\eta}$ is a solution of the system (2.7). When $\eta_3 = 0$ the equations (2.7) imply that $\eta_1 = \eta_2 = 0$, and so there are no non-trivial solutions with $\eta_3 = 0$. We may therefore take $\eta_3 = 1$, and then the equations (2.7) imply that

$$(h - \eta_i)^{k-1} = (h - a_i - \eta_i)^{k-1} \quad (i = 1, 2),$$

whence for $i = 1, 2$, one has $h - \eta_i = \omega_i a_i/(\omega_i - 1)$, where ω_i is some $(k-1)$th root of unity with $\omega_i \neq 1$. Re-substituting into (2.7), we finally deduce that $f(u,v,h)$ is absolutely irreducible unless, for some $(k-1)$th roots of unity ω_i with $\omega_i \neq 1$ $(i = 1, 2)$, the equation

$$(h - a_1)^k + a_1^k(\omega_1 - 1)^{1-k} = (h - a_2)^k + a_2^k(\omega_2 - 1)^{1-k} \tag{2.8}$$

is satisfied. The proof of the lemma is completed on noting that when $a_1 \neq a_2$, for fixed ω_i $(i = 1, 2)$, there are at most $k - 1$ solutions, h, of (2.8).

Lemma 2.6. *Let a_1, a_2, b_1, b_2 be fixed non-zero integers, and define $f(u,v,h;\mathbf{a},\mathbf{b})$ as in the statement of Lemma 2.5. Then when $a_1 \neq a_2$, the number, $M(X;f)$, of solutions of the diophantine equation $f(u,v,h;\mathbf{a},\mathbf{b}) = 0$ with $1 \leq |u|, |v|, |h| \leq X$, satisfies $M(X;f) \ll_{\varepsilon,k} X^{k/(k-1)+\varepsilon}$.*

Proof. Suppose that $a_1 \neq a_2$, and define the set $\mathcal{H}(\mathbf{a},\mathbf{b})$ as in the statement of Lemma 2.5. Denote by $M_1(X;f)$ the number of solutions u, v, h of the equation $f(u,v,h;\mathbf{a},\mathbf{b}) = 0$, with $1 \leq |u|, |v|, |h| \leq X$ and $h \notin \mathcal{H}(\mathbf{a},\mathbf{b})$, and denote by $M_2(X;f)$ the corresponding number of solutions with $h \in \mathcal{H}(\mathbf{a},\mathbf{b})$. Thus

$$M(X;f) = M_1(X;f) + M_2(X;f). \tag{2.9}$$

We start by noting that when $h \notin \mathcal{H}(\mathbf{a},\mathbf{b})$, then by Lemma 2.5 the polynomial $f(u,v,h)$ is absolutely irreducible in $\mathbb{Q}[u,v]$. Moreover, $f(u,v,h)$ has degree $k-1$ as a polynomial in u and v. It therefore follows from Bombieri and Pila [1, Theorem 5], as in the proof of [14, Corollary 2.4.1], that

$$M_1(X;f) \ll_{\varepsilon,k} \sum_{1 \leq |h| \leq X} X^{1/(k-1)+\varepsilon} \ll_{\varepsilon,k} X^{k/(k-1)+\varepsilon}. \tag{2.10}$$

Meanwhile, for each fixed $h \in \mathcal{H}(\mathbf{a},\mathbf{b})$, a trivial estimate shows that there are at most $(k-1)X$ integral solutions u, v of the equation $f(u,v,h) = 0$. Consequently, on noting that the cardinality of $\mathcal{H}$ is at most k^3, we deduce that

$$M_2(X;f) \ll k \sum_{h \in \mathcal{H}} X \ll_k X. \tag{2.11}$$

The lemma follows on combining (2.9), (2.10) and (2.11).

3. The execution of the plan: prelude

When $k \geq 4$, let $N_k(P)$ denote the number of solutions of the system

$$\begin{aligned} u_1^k + u_2^k + u_3^k &= v_1^k + v_2^k + v_3^k, \\ u_1 + u_2 + u_3 &= v_1 + v_2 + v_3, \end{aligned} \tag{3.1}$$

with $1 \leq u_i, v_i \leq P$ $(1 \leq i \leq 3)$, and satisfying the condition that (u_1, u_2, u_3) is not a permutation of (v_1, v_2, v_3). As observed by Greaves [7], the latter condition ensures that $u_i = v_j$ for no i and j. For suppose, to the contrary, that $\mathbf{u}$, $\mathbf{v}$ is a solution counted by $N_k(P)$ with $u_i = v_j$ for some i and j. Since (u_1, u_2, u_3) is not a permutation of (v_1, v_2, v_3), by relabelling variables we may suppose that $u_3 = v_3$ and $u_1 > v_1 \geq v_2 > u_2$. On writing

$$g(u, v) = \sum_{r=0}^{k-1} u^r v^{k-1-r}, \tag{3.2}$$

we deduce from (3.1) that

$$g(u_1, v_1) = \frac{u_1^k - v_1^k}{u_1 - v_1} = \frac{v_2^k - u_2^k}{v_2 - u_2} = g(u_2, v_2). \tag{3.3}$$

But on recalling that $u_2 < u_1$ and $v_2 \leq v_1$, it follows from (3.2) that $g(u_1, v_1) > g(u_2, v_2)$, contradicting (3.3), and establishing the above assertion.

We will find it useful to estimate at this stage the number, $N_k^*(P)$, of solutions $\mathbf{u}$, $\mathbf{v}$ counted by $N_k(P)$ in which $u_3 = v_1 + v_2$. For any such solution it follows from (3.1) that in addition one has $v_3 = u_1 + u_2$. Then by substituting into (3.1), we deduce that $N_k^*(P)$ is bounded above by the number of solutions of the equation

$$(u_1 + u_2)^k - u_1^k - u_2^k = (v_1 + v_2)^k - v_1^k - v_2^k$$

with $1 \leq u_i, v_i \leq P$ $(i = 1, 2)$. Since the polynomial $(x + y)^k - x^k - y^k$ is divisible by both x and y, it follows by using standard estimates for the divisor function that for any fixed v_1 and v_2, there are $O_{\varepsilon,k}(P^\varepsilon)$ possible choices for u_1 and u_2. Consequently,

$$N_k^*(P) \ll_{\varepsilon,k} P^{2+\varepsilon}. \tag{3.4}$$

For each solution $\mathbf{u}$, $\mathbf{v}$ counted by $N_k(P) - N_k^*(P)$, we define the integers x, y, z, w and h by

$$x = u_1 - v_3, \quad y = u_2 - v_3, \quad z = v_1 - u_3, \quad w = v_2 - u_3 \quad \text{and} \quad h = v_1 + v_2. \tag{3.5}$$

We note that the definitions (3.5), together with the linear equation of (3.1), imply that

$$2u_1 = h + z + w - 2y, \quad 2u_2 = h + z + w - 2x, \quad 2u_3 = h - z - w, \tag{3.6}$$

$$2v_1 = h + z - w, \quad 2v_2 = h - z + w, \quad 2v_3 = h + z + w - 2x - 2y. \tag{3.7}$$

Since $u_i = v_j$ for no i and j, and $u_2 + u_3 > 0$, it follows from (3.6), (3.7) and (3.1) that

$$xyzw \neq 0, \quad x \neq z, \quad x \neq w, \quad x \neq h. \tag{3.8}$$

Moreover, on recalling that for solutions counted by $N_k(P) - N_k^*(P)$ one has $u_3 \neq v_1 + v_2$, it follows from (3.6) and (3.7) that

$$h + z + w \neq 0. \tag{3.9}$$

Write

$$F(\alpha, \beta, \gamma) = (\gamma + \alpha - \beta)^k + (\gamma - \alpha + \beta)^k - (\gamma - \alpha - \beta)^k. \tag{3.10}$$

Then we deduce from (3.1), (3.4), (3.6) and (3.7) that

$$N_k(P) \leq M_k(2P) + N_k^*(P) = M_k(2P) + O_{\varepsilon,k}(P^{2+\varepsilon}), \tag{3.11}$$

where $M_k(Q)$ denotes the number of solutions of the equation

$$F(x, y, h + z + w - x - y) = F(z, w, h), \tag{3.12}$$

with the variables satisfying (3.8), (3.9) and

$$1 \leq |x|, |y|, |z|, |w|, |h| \leq Q. \tag{3.13}$$

The argument which yields our theorem takes on a different character according to whether k is odd or even. We treat these respective cases in two lemmata in §§4 and 5 below. Before embarking on the statement and proof of these lemmata, we note a useful property of the polynomial F. An application of the multinomial theorem to (3.10) yields

$$(\gamma + \alpha + \beta)^k - F(\alpha, \beta, \gamma) = \sum_{\substack{r+s+t=k \\ r\geq 0, s\geq 0, t\geq 0}} \frac{k!}{r!s!t!} \gamma^r \alpha^s \beta^t \varepsilon_{rst},$$

where

$$\varepsilon_{rst} = 1 + (-1)^{s+t} - (-1)^s - (-1)^t.$$

On noting that ε_{rst} is zero unless s and t are both odd, and noting that in consequence r must have the same parity as k, we deduce that when k is even,

$$(\gamma + \alpha + \beta)^k - F(\alpha, \beta, \gamma) = \alpha\beta\Psi_k(\alpha, \beta, \gamma), \tag{3.14}$$

where

$$\Psi_k(\alpha, \beta, \gamma) = 4 \sum_{\substack{r+s+t=(k-2)/2 \\ r\geq 0, s\geq 0, t\geq 0}} \frac{k!}{(2r)!(2s+1)!(2t+1)!} \gamma^{2r} \alpha^{2s} \beta^{2t}, \tag{3.15}$$

and when k is odd,

$$(\gamma + \alpha + \beta)^k - F(\alpha, \beta, \gamma) = \alpha\beta\gamma\Phi_k(\alpha, \beta, \gamma), \tag{3.16}$$

where

$$\Phi_k(\alpha, \beta, \gamma) = 4 \sum_{\substack{r+s+t=(k-3)/2 \\ r\geq 0, s\geq 0, t\geq 0}} \frac{k!}{(2r+1)!(2s+1)!(2t+1)!} \gamma^{2r} \alpha^{2s} \beta^{2t}. \tag{3.17}$$

We note for future reference that for real values of α, β and γ, the polynomial $\Psi_k(\alpha, \beta, \gamma)$ is zero if and only if $\alpha = \beta = \gamma = 0$, and likewise for $\Phi_k(\alpha, \beta, \gamma)$.

4. Execution of the plan: the odd case

When k is odd our theorem follows from (3.11) in combination with Lemma 4.1 below. Throughout, implicit constants in the notations of Landau and Vinogradov will depend at most on k, and the positive number ε.

Lemma 4.1. *Let k be an odd integer with $k \geq 5$. Then*

$$M_k(Q) \ll Q^{\frac{8}{3}+\frac{1}{k-1}+\varepsilon}.$$

Proof. For the sake of convenience we write

$$\Omega_k(x,y,z,w,h) = y(h+z+w-x-y)\Phi_k(x,y,h+z+w-x-y). \tag{4.1}$$

Then on recalling (3.12), we find from (3.16) that $M_k(Q)$ is equal to the number of solutions of the equation

$$x\Omega_k(x,y,z,w,h) = zwh\Phi_k(z,w,h), \tag{4.2}$$

with x, y, z, w, h satisfying (3.8), (3.9) and (3.13).

Since k is odd, it follows from (3.17) that $\Phi_k(\alpha,\beta,\gamma)$ is a symmetric function of the variables α, β, γ. Moreover it follows immediately from (4.1) that for each x and y the polynomial $\Omega_k(x,y,\alpha,\beta,\gamma)$ is also symmetric in α, β, γ. Thus if x, y, z, w, h is any solution of (4.2) counted by $M_k(Q)$, we may relabel z, w and h so that the pairwise highest common factors with x satisfy the condition

$$(z,x) \geq \max\{(w,x),(h,x)\} \quad \text{and} \quad \left(w, \frac{x}{(z,x)}\right) \geq \left(h, \frac{x}{(z,x)}\right). \tag{4.3}$$

Write $d=(x,z)$, $e=(x/d,w)$, $f=(x/(de),h)$, and put

$$x_1 = x/(def), \quad z_1 = z/d, \quad w_1 = w/e, \quad h_1 = h/f. \tag{4.4}$$

Then $(x_1, z_1w_1h_1) = 1$, and from (4.3) one has $d \geq e \geq f$. On substituting into (4.2) we obtain

$$x_1\Omega_k(defx_1, y, dz_1, ew_1, fh_1) = z_1w_1h_1\Phi_k(dz_1, ew_1, fh_1). \tag{4.5}$$

We now relate $M_k(Q)$ to the number of solutions of the equation (4.5) satisfying various conditions. Let $S_0(d,e,f)$ denote the number of solutions (x_1,y,z_1,w_1,h_1) of the equation (4.5) with

$$1 \leq |x_1| \leq Q/(def), \tag{4.6}$$

$$1 \leq |y| \leq Q, \tag{4.7}$$

$$1 \leq |z_1| \leq Q/d, \quad (x_1,z_1)=1, \quad z_1 \neq efx_1, \tag{4.8}$$

$$1 \leq |w_1| \leq Q/e, \quad (x_1,w_1)=1, \tag{4.9}$$

$$1 \le |h_1| \le Q/f, \quad (x_1, h_1) = 1 \tag{4.10}$$

and

$$dz_1 + ew_1 + fh_1 \ne 0. \tag{4.11}$$

Then it follows from the above discussion that

$$M_k(Q) \ll \sum_{\substack{def \le Q \\ d \ge e \ge f \ge 1}} S_0(d, e, f). \tag{4.12}$$

We first dispose of the contribution of $S_0(d, e, f)$ to $M_k(Q)$ when $d > Q^{1/3}$. From (3.10), (3.16) and (4.1) we find that the equation (4.5) implies that

$$\begin{aligned}(g - 2defx_1)^k &+ (g - 2y)^k - (g - 2defx_1 - 2y)^k \\ &= (g - 2dz_1)^k + (g - 2ew_1)^k - (g - 2dz_1 - 2ew_1)^k,\end{aligned} \tag{4.13}$$

where $g = dz_1 + ew_1 + fh_1$. Consequently $S_0(d, e, f)$ is bounded above by the number of solutions of the equation (4.13) with (4.6)-(4.9) and $1 \le |g| \le 3Q$. For each fixed choice of x_1 and z_1, we solve the equation (4.13) for y, w_1 and g. On noting that (4.8) implies that $efx_1 \ne z_1$, we may appeal to Lemma 2.6, therefore, to deduce that the number of possible choices for y, w_1 and g is $O(Q^{k/(k-1)+\varepsilon})$. Consequently,

$$\begin{aligned}S_0(d, e, f) &\ll \sum_{1 \le |x_1| \le Q/(def)} \sum_{1 \le |z_1| \le Q/d} Q^{k/(k-1)+\varepsilon} \\ &\ll (d^2ef)^{-1} Q^{3 + \frac{1}{k-1} + \varepsilon}.\end{aligned}$$

Thus we conclude that

$$\begin{aligned}M_k(Q) - \sum_{\substack{def \le Q \\ 1 \le f \le e \le d \le Q^{1/3}}} S_0(d, e, f) &\ll \sum_{d > Q^{1/3}} \sum_{1 \le ef \le Q} (d^2ef)^{-1} Q^{3 + \frac{1}{k-1} + \varepsilon} \\ &\ll Q^{3 + \frac{1}{k-1} + 2\varepsilon} \sum_{d > Q^{1/3}} d^{-2} \\ &\ll Q^{\frac{8}{3} + \frac{1}{k-1} + 2\varepsilon}.\end{aligned} \tag{4.14}$$

We now consider the contribution of $S_0(d, e, f)$ to $M_k(Q)$ when $d \le Q^{1/3}$. Write

$$\mathcal{X} = \max\left\{ \frac{15Q^{2/3}}{(df)^{1/3} e^{4/3}}, \frac{2(Qd)^{1/3}}{(ef)^{2/3}}, \frac{3Q^{1/2}}{e^{1/2} f} \right\}. \tag{4.15}$$

We denote by $S_1(d, e, f)$ the number of solutions (x_1, y, z_1, w_1, h_1) counted by $S_0(d, e, f)$ in which

$$1 \le |x_1| \le \mathcal{X}, \tag{4.16}$$

and let $S_2(d,e,f)$ denote the corresponding number of solutions in which

$$\mathcal{X} < |x_1| \le Q/(def). \tag{4.17}$$

Then it follows from (4.14) that

$$M_k(Q) \ll Q^{\frac{8}{3}+\frac{1}{k-1}+2\varepsilon} + \sum_{\substack{def\le Q\\ 1\le f\le e\le d\le Q^{1/3}}} \left(S_1(d,e,f)+S_2(d,e,f)\right). \tag{4.18}$$

Next we estimate $S_1(d,e,f)$, noting that given x_1 and z_1 we may proceed, as in the argument leading to (4.14), to solve the equation (4.13) for y, w_1 and g. Thus, in view of (4.16), we arrive at the conclusion

$$S_1(d,e,f) \ll \sum_{1\le |x_1|\le \mathcal{X}} \sum_{1\le |z_1|\le Q/d} Q^{k/(k-1)+\varepsilon} \ll d^{-1}\mathcal{X}Q^{2+\frac{1}{k-1}+\varepsilon}.$$

Then by (4.15) we obtain

$$\sum_{1\le f\le e\le d\le Q^{1/3}} S_1(d,e,f) \ll Q^{2+\frac{1}{k-1}+\varepsilon}\mathcal{U}(Q),$$

where

$$\begin{aligned}\mathcal{U}(Q) &= \sum_{1\le f\le e\le d\le Q^{1/3}} \left(\frac{Q^{2/3}}{(de)^{4/3}f^{1/3}} + \frac{Q^{1/3}}{(def)^{2/3}} + \frac{Q^{1/2}}{dfe^{1/2}}\right)\\ &\ll \sum_{1\le e\le d\le Q^{1/3}} \left(\frac{Q^{2/3}}{d^{4/3}e^{2/3}} + \frac{Q^{1/3}}{d^{2/3}e^{1/3}} + \frac{Q^{1/2+\varepsilon}}{de^{1/2}}\right)\\ &\ll \sum_{1\le d\le Q^{1/3}} \left(Q^{2/3}d^{-1} + Q^{1/3} + Q^{1/2+\varepsilon}d^{-1/2}\right) \ll Q^{2/3+\varepsilon}.\end{aligned}$$

Consequently,

$$\sum_{1\le f\le e\le d\le Q^{1/3}} S_1(d,e,f) \ll Q^{\frac{8}{3}+\frac{1}{k-1}+2\varepsilon}. \tag{4.19}$$

Next consider the solutions (x_1,y,z_1,w_1,h_1) counted by $S_2(d,e,f)$. By Lemma 2.2, for each 4-tuple (x_1,z_1,w_1,h_1) satisfying (4.6) and (4.8)-(4.10), there exists a real number X with

$$1 \le X \le 2Q^{4/3}(def)^{-2/3}, \tag{4.20}$$

and linearly independent integral 4-tuples $\mathbf{a}$ and $\mathbf{b}$ with

$$|a_i| \le \beta_i Q^{-1}X \quad \text{and} \quad |b_i| \le 5\beta_i(def)^{-1}QX^{-1/2} \quad (1\le i\le 4), \tag{4.21}$$

where

$$\beta_1 = f, \quad \beta_2 = e, \quad \beta_3 = d, \quad \beta_4 = def, \tag{4.22}$$

and satisfying the equations

$$\alpha_1 h_1 + \alpha_2 w_1 = \alpha_3 z_1 + \alpha_4 x_1 \quad (\alpha = a, b). \tag{4.23}$$

We note that when $a_i \neq 0$, the condition (4.21) implies that necessarily,

$$X \geq Q/\beta_i. \tag{4.24}$$

Solving the equations (4.23) for h_1 and w_1, we obtain

$$\lambda h_1 = \mu_1 z_1 + \nu_1 x_1, \quad \lambda w_1 = \mu_2 z_1 - \nu_2 x_1, \tag{4.25}$$

where

$$\lambda = a_1 b_2 - a_2 b_1, \quad \mu_1 = a_3 b_2 - a_2 b_3, \quad \mu_2 = a_1 b_3 - a_3 b_1, \tag{4.26}$$

$$\nu_1 = a_4 b_2 - a_2 b_4, \quad \nu_2 = a_4 b_1 - a_1 b_4. \tag{4.27}$$

Further, by Lemma 2.1, for each triple (x_1, z_1, h_1) satisfying (4.6), (4.8) and (4.10), there exist integers c_1, c_3, c_4 with $(c_1, c_3, c_4) = 1$ and

$$|c_i| \leq 3\beta_i Q^{1/2}(d^2 e f^2)^{-1/2}, \tag{4.28}$$

and satisfying the equation

$$c_1 h_1 = c_3 z_1 - c_4 x_1. \tag{4.29}$$

We claim that for each solution (x_1, y, z_1, w_1, h_1) counted by $S_2(d, e, f)$, there exist integral vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ satisfying the conditions of the previous paragraph, and in addition satisfying either (i) $a_1 \neq 0$ and $\lambda\mu_1\mu_2 \neq 0$ or (ii) $a_1 = 0$ and $a_2 a_3 c_1 c_3 \neq 0$.

In order to establish this proposition we first observe that if $c_1 c_3 = 0$, then (4.29) together with (4.8) and (4.10) implies that we have either $c_3 z_1 = c_4 x_1$ with $(x_1, z_1) = (c_3, c_4) = 1$, or else $c_1 h_1 = -c_4 x_1$ with $(x_1, h_1) = (c_1, c_4) = 1$. Consequently, either $x_1 | c_3$ or $x_1 | c_1$, whence by (4.28) and (4.22), we have

$$|x_1| \leq |c_3| \leq 3Q^{1/2}(ef^2)^{-1/2} \quad \text{or} \quad |x_1| \leq |c_1| \leq 3Q^{1/2}(d^2 e)^{-1/2}.$$

But (4.15) together with the inequality $f \leq d$ implies that the latter inequalities contradict (4.17), so that necessarily $c_1 c_3 \neq 0$. Suppose next that $a_1 = 0$ and $a_2 a_3 = 0$. It follows from (4.23) that if two of the a_i are zero, then the two remaining a_j must be non-zero. Then by (4.8), (4.9) and (4.23) we have either $a_3 z_1 = -a_4 x_1$ with $(x_1, z_1) = 1$ and $a_3 a_4 \neq 0$, or else $a_2 w_1 = a_4 x_1$ with $(x_1, w_1) = 1$ and $a_2 a_4 \neq 0$. Consequently, either $x_1 | a_3$ or $x_1 | a_2$, whence by (4.21) and (4.22), we have

$$|x_1| \leq |a_3| \leq 2(Qd)^{1/3}(ef)^{-2/3} \quad \text{or} \quad |x_1| \leq |a_2| \leq 2(Qe)^{1/3}(df)^{-2/3}.$$

But (4.15) together with the inequality $e \le d$ implies that the latter inequalities contradict (4.17), so that necessarily $a_2a_3 \neq 0$. Thus whenever $a_1 = 0$, one has $a_2a_3c_1c_3 \neq 0$.

Next suppose that $a_1 \neq 0$. Consider first the case $\lambda \neq 0$ and $\mu_1\mu_2 = 0$. Then from (4.25), (4.9) and (4.10) we have either $\lambda h_1 = \nu_1 x_1$ with $(x_1, h_1) = 1$, or else $\lambda w_1 = -\nu_2 x_1$ with $(x_1, w_1) = 1$. Consequently, in either case one has $x_1|\lambda$, whence by (4.20)-(4.22) and (4.26),

$$|x_1| \le |\lambda| \le 10X^{1/2}d^{-1} \le 15Q^{2/3}d^{-4/3}(ef)^{-1/3},$$

and since $d \ge e$, on recalling (4.15), we obtain a contradiction to (4.17). Meanwhile, if $\lambda = 0$ and $\mu_2 \neq 0$, then by (4.8) and (4.25) we have $\mu_2 z_1 = \nu_2 x_1$ with $(x_1, z_1) = 1$. Thus $x_1|\mu_2$, whence by (4.20)-(4.22) and (4.26),

$$|x_1| \le |\mu_2| \le 10X^{1/2}e^{-1} \le 15Q^{2/3}e^{-4/3}(df)^{-1/3},$$

which in view of (4.15) contradicts (4.17). Then we are left to deal with the case $\lambda = \mu_2 = 0$. In this latter case, (4.25) and (4.6) imply that $\nu_2 = 0$. But (4.26) and (4.27) imply that $a_1\mathbf{b} = b_1\mathbf{a}$, so that since $a_1 \neq 0$ we obtain a contradiction to the linear independence of $\mathbf{a}$ and $\mathbf{b}$. This completes the proof of our assertion.

Let $S_3^{(1)}(d, e, f; Y)$ denote the number of solutions of the equation (4.5) satisfying (4.6)-(4.11), such that there exists a real number X with $Y \le X \le 2Y$, and linearly independent 4-tuples $\mathbf{a}$ and $\mathbf{b}$ satisfying (4.21)-(4.23) and, in the notation defined by (4.26), satisfying the additional conditions $a_1 \neq 0$ and $\lambda\mu_1\mu_2 \neq 0$. Also, let $S_3^{(2)}(d, e, f; Y)$ denote the corresponding number of solutions with the property that there exist triples (a_2, a_3, a_4) and (c_1, c_3, c_4) satisfying (4.29), (4.21),

$$a_2 w_1 = a_3 z_1 + a_4 x_1, \tag{4.30}$$

and $a_2a_3c_1c_3 \neq 0$. Then by considering dyadic intervals for Y, and recalling (4.24), one obtains

$$S_2(d, e, f) \ll (\mathcal{S}_1 + \mathcal{S}_2)\log(2Q), \tag{4.31}$$

where

$$\mathcal{S}_1 = \max_{Q/f \le Y \le Q^{4/3}(def)^{-2/3}} S_3^{(1)}(d, e, f; Y)$$

and

$$\mathcal{S}_2 = \max_{Q/e \le Y \le Q^{4/3}(def)^{-2/3}} S_3^{(2)}(d, e, f; Y).$$

Let $S_4^{(1)}(d, e, f; \mathbf{a}, \mathbf{b})$ denote the number of solutions x_1, y, z_1, w_1, h_1 of the equation (4.5) satisfying (4.6)-(4.11) and (4.25)-(4.27), and let $S_4^{(2)}(d, e, f; \mathbf{a}, \mathbf{c})$ denote the corresponding number of solutions satisfying instead (4.6)-(4.11) and (4.29), (4.30). Then

$$S_3^{(1)}(d, e, f; Y) \le \sum_{\mathbf{a},\mathbf{b}} S_4^{(1)}(d, e, f; \mathbf{a}, \mathbf{b}), \tag{4.32}$$

where the summation is over $\mathbf{a}$ and $\mathbf{b}$ with

$$|a_i| \le 2\beta_i Q^{-1}Y \quad \text{and} \quad |b_i| \le 5\beta_i (def)^{-1} QY^{-1/2} \quad (1 \le i \le 4), \tag{4.33}$$

and, in the notation defined by (4.26), satisfying $a_1 \neq 0$ and $\lambda\mu_1\mu_2 \neq 0$. Also,

$$S_3^{(2)}(d,e,f;Y) \le \sum_{\mathbf{a},\mathbf{c}} S_4^{(2)}(d,e,f;\mathbf{a},\mathbf{c}), \tag{4.34}$$

where the summation is over $\mathbf{a}$ and $\mathbf{c}$ with

$$|a_i| \le 2\beta_i Q^{-1}Y \quad (i=2,3,4) \quad \text{and} \quad |c_i| \le 3\beta_i Q^{1/2}(d^2ef^2)^{-1/2} \quad (i=1,3,4), \tag{4.35}$$

and satisfying $a_2a_3c_1c_3 \neq 0$.

We first estimate $S_3^{(1)}$. We substitute from (4.25) for h_1 and w_1 into (4.5) to deduce that $S_4^{(1)}(d,e,f;\mathbf{a},\mathbf{b})$ is bounded above by $S_5^{(1)}(d,e,f;\mathbf{a},\mathbf{b})$, where $S_5^{(1)}$ denotes the number of solutions of the equation

$$\begin{aligned} x_1\Omega_k\left(\lambda defx_1, \lambda y, \lambda dz_1, e(\mu_2z_1-\nu_2x_1), f(\mu_1z_1+\nu_1x_1)\right) \\ = z_1(\mu_2z_1-\nu_2x_1)(\mu_1z_1+\nu_1x_1)\Phi_k\left(\lambda dz_1, e(\mu_2z_1-\nu_2x_1), f(\mu_1z_1+\nu_1x_1)\right), \end{aligned} \tag{4.36}$$

with x_1, y, z_1 satisfying (4.6)-(4.8). Observe that for each solution (x_1,y,z_1) counted by $S_5^{(1)}$, the equation (4.36) implies that

$$\mu_1\mu_2z_1^3\Phi_k(\lambda dz_1, e\mu_2z_1, f\mu_1z_1) \equiv 0 \pmod{x_1}.$$

Then it follows from the coprimality condition of (4.8) together with the homogeneity of Φ_k that

$$x_1 | \mu_1\mu_2\Phi_k(\lambda d, e\mu_2, f\mu_1).$$

Furthermore, since none of λ, μ_1, μ_2 are zero, we have $\mu_1\mu_2\Phi_k(d\lambda, e\mu_2, f\mu_1) \neq 0$, so that by using standard estimates for the divisor function, there are at most $O(Q^\varepsilon)$ possible choices for x_1. Fixing any one such choice, and recalling the definitions of Ω_k and Φ_k, the equation (4.36) takes the shape

$$p(z_1) + yz_1q(y,z_1) + r(y) = 0, \tag{4.37}$$

where $p(z_1)$ is a polynomial of degree k, the polynomial $q(y,z_1)$ has total degree $k-3$, and $r(y)$ has degree $k-1$. It follows from Lemma 2.4 that the number of possible choices for y and z_1 satisfying (4.7) and (4.8) is $O(Q^{1/k+\varepsilon})$. Thus

$$S_4^{(1)}(d,e,f;\mathbf{a},\mathbf{b}) \le S_5^{(1)}(d,e,f;\mathbf{a},\mathbf{b}) \ll Q^{1/k+2\varepsilon},$$

and hence by (4.32),

$$S_3^{(1)}(d,e,f;Y) \ll Q^{1/k+2\varepsilon} \sum_{\mathbf{a},\mathbf{b}} 1. \tag{4.38}$$

Next we note from (4.31) that we have $Q/f \le Y \le Q^{4/3}(def)^{-2/3}$, so that by (4.22), for $1 \le i \le 4$ we have

$$\beta_i Q^{-1} Y \gg \beta_i/f \gg 1,$$

and on recalling the inequality $1 \le f \le e \le d \le Q^{1/3}$, for $i = 2, 3, 4$ we have

$$\beta_i (def)^{-1} Q Y^{-1/2} \gg \beta_i Q^{1/3}(def)^{-2/3} \gg (Qe)^{1/3}(df)^{-2/3} \gg Q^{1/3}(d^2 f)^{-1/3} \gg 1.$$

Consequently, by (4.33) in combination with the above inequality on Y,

$$\begin{aligned} \sum_{\mathbf{a},\mathbf{b}} 1 &\ll \left(1 + \beta_1 (def)^{-1} Q Y^{-1/2}\right) \beta_1 Q^{-1} Y \left(\prod_{i=2,3,4} \beta_i^2 Y^{1/2} (def)^{-1} \right) \\ &\ll \left(1 + Q^{1/3}(de)^{-2/3} f^{1/3}\right) Q^{7/3}(de)^{-2/3} f^{-5/3}. \end{aligned}$$

Thus by (4.38),

$$S_3^{(1)}(d,e,f;Y) \ll Q^{\frac{8}{3}+\frac{1}{k}+2\varepsilon} \left(Q^{-1/3}(de)^{-2/3} f^{-5/3} + (def)^{-4/3}\right). \tag{4.39}$$

Next we estimate $S_3^{(2)}$. We substitute from (4.29) and (4.30) for h_1 and w_1 into (4.5) to deduce that $S_4^{(2)}(d,e,f;\mathbf{a},\mathbf{c})$ is bounded above by $S_5^{(2)}(d,e,f;\mathbf{a},\mathbf{c})$, where $S_5^{(2)}$ denotes the number of solutions of the equation

$$x_1 \Omega_k = a_2 c_1 z_1 (a_3 z_1 + a_4 x_1)(c_3 z_1 - c_4 x_1) \Phi_k, \tag{4.40}$$

with x_1, y, z_1 satisfying (4.6)-(4.8), where we have written

$$\Omega_k = \Omega_k \left(a_2 c_1 def x_1, a_2 c_1 y, a_2 c_1 d z_1, c_1 e (a_3 z_1 + a_4 x_1), a_2 f (c_3 z_1 - c_4 x_1)\right),$$

and

$$\Phi_k = \Phi_k \left(a_2 c_1 d z_1, c_1 e (a_3 z_1 + a_4 x_1), a_2 f (c_3 z_1 - c_4 x_1)\right).$$

Observe that by an argument paralleling the treatment of $S_4^{(1)}$, we have for each solution (x_1, y, z_1) counted by $S_5^{(2)}$, that

$$x_1 | a_2 a_3 c_1 c_3 \Phi_k (a_2 c_1 d, a_3 c_1 e, a_2 c_3 f).$$

Furthermore, since none of a_2, a_3, c_1, c_3 are zero, we have

$$a_2 a_3 c_1 c_3 \Phi_k (a_2 c_1 d, a_3 c_1 e, a_2 c_3 f) \neq 0,$$

so that by using standard estimates for the divisor function, there are at most $O(Q^\varepsilon)$ possible choices for x_1. Fixing any one such choice, we deduce as in the treatment of $S_4^{(1)}$ that the number of possible choices for y and z_1 satisfying (4.7) and (4.8) is $O(Q^{1/k+\varepsilon})$. Thus

$$S_4^{(2)}(d,e,f;\mathbf{a},\mathbf{c}) \ll S_5^{(2)}(d,e,f;\mathbf{a},\mathbf{c}) \ll Q^{1/k+2\varepsilon},$$

and hence by (4.34),

$$S_3^{(2)}(d,e,f;Y) \ll Q^{1/k+2\varepsilon} \sum_{\mathbf{a},\mathbf{c}} 1. \tag{4.41}$$

Next we note from (4.31) that we have $Q/e \le Y \le Q^{4/3}(def)^{-2/3}$, so that by (4.22), for $i = 2,3,4$ we have

$$\beta_i Q^{-1} Y \gg \beta_i/e \gg 1,$$

and on recalling the inequality $1 \le f \le e \le d \le Q^{1/3}$, for $1 \le i \le 4$ we have

$$\beta_i Q^{1/2}(d^2ef^2)^{-1/2} \gg Q^{1/2}(d^2e)^{-1/2} \gg 1.$$

Consequently, by (4.35) in combination with the above inequality on Y,

$$\begin{aligned}\sum_{\mathbf{a},\mathbf{c}} 1 &\ll \beta_1\beta_2 Q^{-1/2} Y (d^2ef^2)^{-1/2} \left(\prod_{i=3,4} \beta_i^2 Q^{-1/2} Y (d^2ef^2)^{-1/2} \right) \\ &\ll Q^{5/2}(de^{1/2}f^2)^{-1}.\end{aligned}$$

Thus by (4.41),

$$S_3^{(2)}(d,e,f;Y) \ll Q^{\frac{5}{2}+\frac{1}{k}+2\varepsilon}(de^{1/2}f^2)^{-1}. \tag{4.42}$$

Finally, combining (4.39), (4.42), (4.31), we find that

$$\sum_{1\le f\le e\le d\le Q^{1/3}} S_2(d,e,f) \ll Q^{\frac{8}{3}+\frac{1}{k}+3\varepsilon}\mathcal{V}(Q), \tag{4.43}$$

where

$$\mathcal{V}(Q) = \sum_{1\le f\le e\le d\le Q^{1/3}} \left(Q^{-1/3}(de)^{-2/3}f^{-5/3} + (def)^{-4/3} + Q^{-1/6}(de^{1/2}f^2)^{-1} \right).$$

But

$$\begin{aligned}\mathcal{V}(Q) &\ll \sum_{1\le e\le d\le Q^{1/3}} \left(Q^{-1/3}(de)^{-2/3} + (de)^{-4/3} + Q^{-1/6}(de^{1/2})^{-1} \right) \\ &\ll \sum_{1\le d\le Q^{1/3}} \left(Q^{-1/3}d^{-1/3} + d^{-4/3} + Q^{-1/6}d^{-1/2} \right) \ll 1.\end{aligned}$$

Thus, on recalling (4.18), (4.19) and (4.43), we deduce that

$$M_k(Q) \ll Q^{\frac{8}{3}+\frac{1}{k-1}+2\varepsilon},$$

and this completes the proof of the lemma.

5. Execution of the plan: the even case

We consider the case in which k is even in Lemma 5.1 below. By combining Lemmata 4.1 and 5.1 with (3.11), the proof of our theorem is completed.

Lemma 5.1. *Let k be an even integer with $k \geq 4$. Then*

$$M_k(Q) \ll_{\varepsilon,k} Q^{\frac{8}{3}+\frac{1}{k-1}+\varepsilon}.$$

Proof. For the sake of convenience we write

$$\Upsilon_k(x,y,z,w,h) = y\Psi_k(x,y,h+z+w-x-y). \tag{5.1}$$

Then on recalling (3.12), we find from (3.14) that $M_k(Q)$ is equal to the number of solutions of the equation

$$x\Upsilon_k(x,y,z,w,h) = zw\Psi_k(z,w,h), \tag{5.2}$$

with x, y, z, w, h satisfying (3.8), (3.9) and (3.13).

Since k is even, it follows from (3.15) that $\Psi_k(\alpha,\beta,\gamma)$ is symmetric in α and β. Moreover it follows immediately from (5.1) that for each x and y the polynomial $\Upsilon_k(x,y,\alpha,\beta,\gamma)$ is symmetric in α and β. Thus if x, y, z, w, h is any solution of (5.2) counted by $M_k(Q)$, we may relabel z and w so that the pairwise highest common factors with x satisfy the condition

$$(z,x) \geq (w,x). \tag{5.3}$$

Write $d = (x,z)$ and $e = (x/d,w)$, and put

$$x_1 = x/(de), \quad z_1 = z/d \quad \text{and} \quad w_1 = w/e. \tag{5.4}$$

Then $(x_1, z_1w_1) = 1$, and from (5.3) one has $d \geq e$. On substituting into (5.2) we obtain

$$x_1\Upsilon_k(dex_1,y,dz_1,ew_1,h) = z_1w_1\Psi_k(dz_1,ew_1,h). \tag{5.5}$$

We now estimate $M_k(Q)$ using an argument strikingly similar, though simpler, than that used in the proof of Lemma 4.1. In order to curtail our deliberations, we adopt the convention throughout the remainder of the proof of this lemma that $f = 1$ and $h_1 = h$. Let $S_0(d,e)$ denote the number of solutions (x_1,y,z_1,w_1,h) of the equation (5.5) satisfying the conditions (4.6)-(4.11), let $S_1(d,e)$ denote the corresponding number of solutions subject to the additional condition (4.16), and let $S_2(d,e)$ denote the corresponding number of solutions subject instead to the additional condition (4.17). Then it follows from the above discussion that

$$M_k(Q) \ll \sum_{\substack{de\leq Q \\ d>Q^{1/3}}} S_0(d,e) + \sum_{\substack{de\leq Q \\ 1\leq e\leq d\leq Q^{1/3}}} \left(S_1(d,e)+S_2(d,e)\right). \tag{5.6}$$

We first observe that from (3.10), (3.14) and (5.1), the equation (5.5) implies that the equation (4.13) is satisfied. Thus the arguments leading to (4.14) and (4.19) remain valid, and we deduce that

$$\sum_{\substack{de\le Q\\ d>Q^{1/3}}} S_0(d,e)+\sum_{\substack{de\le Q\\ 1\le e\le d\le Q^{1/3}}} S_1(d,e)\ll Q^{\frac{8}{3}+\frac{1}{k-1}+\varepsilon}. \tag{5.7}$$

We estimate $S_2(d,e)$ when $1\le e\le d\le Q^{1/3}$ as in the argument used to estimate $S_2(d,e,f)$ in the proof of Lemma 4.1. Let $U_1^{(1)}(d,e;\mathbf{a},\mathbf{b})$ denote the number of solutions (x_1,y,z_1,w_1,h) of the equation (5.5) satisfying (4.6)-(4.11) and (4.25)-(4.27), and let $U_1^{(2)}(d,e;\mathbf{a},\mathbf{c})$ denote the corresponding number of solutions satisfying instead (4.6)-(4.11) and (4.29), (4.30). Then it follows, as in the argument leading to (4.31), (4.32) and (4.34), that

$$S_2(d,e)\ll(\mathcal{T}_1+\mathcal{T}_2)\log(2Q), \tag{5.8}$$

where

$$\mathcal{T}_1=\max_{Q\le Y\le Q^{4/3}(de)^{-2/3}}\sum_{\mathbf{a},\mathbf{b}}U_1^{(1)}(d,e;\mathbf{a},\mathbf{b}),$$

$$\mathcal{T}_2=\max_{Q/e\le Y\le Q^{4/3}(de)^{-2/3}}\sum_{\mathbf{a},\mathbf{c}}U_1^{(2)}(d,e;\mathbf{a},\mathbf{c}),$$

and the respective summations are over $\mathbf{a}$ and $\mathbf{b}$ satisfying (4.33) with (4.22), and in the notation defined by (4.26), subject to $a_1\ne 0$ and $\lambda\mu_1\mu_2\ne 0$, and over $\mathbf{a}$ and $\mathbf{c}$ satisfying (4.35) with (4.22), and subject to $a_2a_3c_1c_3\ne 0$.

On substituting from (4.25) for h and w_1 into (5.5), we deduce that $U_1^{(1)}(d,e;\mathbf{a},\mathbf{b})$ is bounded above by $U_2^{(1)}=U_2^{(1)}(d,e;\mathbf{a},\mathbf{b})$, where $U_2^{(1)}$ denotes the number of solutions of the equation

$$\begin{aligned}x_1\Upsilon_k\left(\lambda dex_1,\lambda y,\lambda dz_1,e(\mu_2z_1-\nu_2x_1),\mu_1z_1+\nu_1x_1\right)&\\ =z_1(\mu_2z_1-\nu_2x_1)\Psi_k\left(\lambda dz_1,e(\mu_2z_1-\nu_2x_1),\mu_1z_1+\nu_1x_1\right),&\end{aligned} \tag{5.9}$$

with x_1,y,z_1 satisfying (4.6)-(4.8). Observe that for each solution (x_1,y,z_1) counted by $U_2^{(1)}$, the equation (5.9) implies that

$$\mu_2z_1^2\Psi_k(\lambda dz_1,e\mu_2z_1,\mu_1z_1)\equiv 0 \pmod{x_1}.$$

Then it follows from the coprimality condition of (4.8) together with the homogeneity of Ψ_k that

$$x_1|\mu_2\Psi_k(\lambda d,e\mu_2,\mu_1).$$

Furthermore, since neither λ nor μ_2 are zero, we have $\mu_2\Psi_k(\lambda d,e\mu_2,\mu_1)\ne 0$, so that by using standard estimates for the divisor function, there are at most $O(Q^\varepsilon)$ possible choices for x_1. Fixing any one such choice, and recalling the definitions of

Υ_k and Ψ_k, the equation (5.9) takes the shape (4.37). Thus, as in the argument leading to (4.38), we obtain

$$U_1^{(1)}(d,e;\mathbf{a},\mathbf{b}) \le U_2^{(1)}(d,e;\mathbf{a},\mathbf{b}) \ll Q^{1/k+2\varepsilon},$$

and hence, following the argument leading to (4.39), we obtain

$$\sum_{\mathbf{a},\mathbf{b}} U_1^{(1)}(d,e;\mathbf{a},\mathbf{b}) \ll Q^{\frac{8}{3}+\frac{1}{k}+2\varepsilon}\left(Q^{-1/3}(de)^{-2/3}+(de)^{-4/3}\right). \tag{5.10}$$

Next we estimate $U_1^{(2)}$. We substitute from (4.29) and (4.30) for h and w_1 into (5.5) to deduce that $U_1^{(2)}(d,e;\mathbf{a},\mathbf{c})$ is bounded above by $U_2^{(2)}(d,e;\mathbf{a},\mathbf{c})$, where $U_2^{(2)}$ denotes the number of solutions of the equation

$$x_1\Upsilon_k = c_1z_1(a_3z_1+a_4x_1)\Psi_k, \tag{5.11}$$

with x_1,y,z_1 satisfying (4.6)-(4.8), where we have written

$$\Upsilon_k = \Upsilon_k\left(a_2c_1dex_1, a_2c_1y, a_2c_1dz_1, c_1e(a_3z_1+a_4x_1), a_2(c_3z_1-c_4x_1)\right),$$

and

$$\Psi_k = \Psi_k\left(a_2c_1dz_1, c_1e(a_3z_1+a_4x_1), a_2(c_3z_1-c_4x_1)\right).$$

Observe that by an argument paralleling the treatment of $S_4^{(1)}$, we have for each solution (x_1,y,z_1) counted by $U_2^{(2)}$, that

$$x_1|a_3c_1\Psi_k(a_2c_1d, a_3c_1e, a_2c_3).$$

Furthermore, since none of a_2,a_3,c_1,c_3 are zero, we have

$$a_3c_1\Psi_k(a_2c_1d, a_3c_1e, a_2c_3) \ne 0,$$

so that by using standard estimates for the divisor function, there are at most $O(Q^\varepsilon)$ possible choices for x_1. Fixing any one such choice, and recalling the definitions of Υ_k and Ψ_k, the equation (5.11) takes the shape (4.37). Thus, as in the argument leading to (4.41), we obtain

$$U_1^{(2)}(d,e;\mathbf{a},\mathbf{c}) \le U_2^{(2)}(d,e;\mathbf{a},\mathbf{c}) \ll Q^{1/k+2\varepsilon},$$

and hence, following the argument leading to (4.42), we obtain

$$\sum_{\mathbf{a},\mathbf{c}} U_1^{(2)}(d,e;\mathbf{a},\mathbf{c}) \ll Q^{\frac{5}{2}+\frac{1}{k}+2\varepsilon}\left(de^{1/2}\right)^{-1}. \tag{5.12}$$

Finally, combining (5.10), (5.12), (5.8), we find that

$$\sum_{1\le e\le d\le Q^{1/3}} S_2(d,e) \ll Q^{\frac{8}{3}+\frac{1}{k}+3\varepsilon}\mathcal{V}(Q), \tag{5.13}$$

where

$$\mathcal{V}(Q) = \sum_{1\le e\le d\le Q^{1/3}} \left(Q^{-1/3}(de)^{-2/3}+(de)^{-4/3}+Q^{-1/6}(de^{1/2})^{-1}\right) \ll 1.$$

Thus, on recalling (5.6), (5.7) and (5.13), we deduce that

$$M_k(Q) \ll Q^{\frac{8}{3}+\frac{1}{k-1}+3\varepsilon},$$

and this completes the proof of the lemma.

References

1. E. Bombieri and J. Pila, *The number of integral points on arcs and ovals*, Duke Math. J. **59** (1989), 337–357.
2. E. Bombieri and J. Vaaler, *On Siegel's Lemma*, Invent. Math. **73** (1983), 11–32.
3. A. Bremner, *A geometric approach to equal sums of fifth powers*, J. Number Theory **13** (1981), 337–354.
4. L. E. Dickson, *History of the Theory of Numbers*, vol. 2, Chelsea, New-York, 1952.
5. A. Gloden, *Mehrgradige Gleichungen*, Noordhoff, Groningen, 1944.
6. G. Greaves, *On the representation of a number as the sum of two fourth powers*, Math. Z. **94** (1966), 223–234.
7. G. Greaves, *Some diophantine equations with almost all solutions trivial*, Mathematika (to appear).
8. C. Hooley, *On the representation of a number as a sum of two cubes*, Math. Z. **82** (1963), 259–266.
9. C. Hooley, *On the representation of a number as a sum of two hth powers*, Math. Z. **84** (1964), 126–136.
10. C. Hooley, *On the numbers that are representable as the sum of two cubes*, J. Reine Angew. Math. **314** (1980), 146–173.
11. C. Hooley, *On another sieve method and the numbers that are a sum of two hth powers*, Proc. London Math. Soc. (3) **43** (1981), 73–109.
12. N. N. Rogovskaya, *An asymptotic formula for the number of solutions of a system of equations*, Diophantine approximations, Part II, Moskov. Gos. Univ. Moscow, 1986, pp. 78–84. (Russian)
13. W. M. Schmidt, *Equations over finite fields. An elementary approach. Lecture Notes in Mathematics, Vol. 536*, Springer-Verlag, Berlin-New York, 1976.
14. C. M. Skinner and T. D. Wooley, *Sums of two kth powers*, J. Reine Angew. Math. **462** (1995), 57–68.
15. R. C. Vaughan and T. D. Wooley, *On a certain nonary cubic form and related equations*, Duke Math. J. (to appear).
16. T. D. Wooley, *A note on symmetric diagonal equations*, Number Theory with an emphasis on the Markoff spectrum (A. D. Pollington and W. Moran, ed.), Marcel Dekker, New York, 1993, pp. 317–321.

CMS: Department of Mathematics, Princeton University, Princeton, New Jersey 08544

E-mail address: `cmcls@phoenix.princeton.edu`

TDW: Mathematics Department, University of Michigan, Ann Arbor, Michigan, 48109-1003

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