SUMS AND DIFFERENCES OF FOUR k-TH POWERS

OSCAR MARMON

Abstract. We prove an upper bound for the number of representations of a positive integer N as the sum of four k-th powers of integers of size at most B, using a new version of the Determinant method developed by Heath-Brown, along with recent results by Salberger on the density of integral points on affine surfaces.

More generally we consider representations by any integral diagonal form. The upper bound has the form $O_N(B^{c/\sqrt{k}})$, whereas earlier versions of the Determinant method would produce an exponent for B of order $k^{-1/3}$ in this case.

Furthermore, we prove that the number of representations of a positive integer N as a sum of four k-th powers of non-negative integers is at most $O_\epsilon(N^{1/k+2/k^{3/2}+\epsilon})$ for $k \geq 3$, improving upon bounds by Wisdom.

1. Introduction

In this paper, we shall study the number of representations of a positive integer N using four k-th powers. We consider two different versions of this problem. The main part of the paper concerns solutions to the equation

$$x_1^k \pm x_2^k \pm x_3^k \pm x_4^k = N \tag{1}$$

in integers x_i, positive or negative. Our treatment of this problem is inspired by a recent paper of Heath-Brown [8] where he studies the equation

$$x_1^k \pm x_2^k \pm x_3^k = N. \tag{2}$$

More precisely, he estimates the number of integral solutions to (2), with $\max |x_i| \leq B$, that are not trivial in the sense that $\pm x_i^k = N$ for some i. Assuming that $N \ll B$, Heath-Brown proves that there are $O_k(B^{10/k})$ such solutions for $k \geq 3$.

The method used by Heath-Brown is a new approach to the determinant method of Bombieri and Pila [1]. Rather than counting integral points on the affine surface defined by (2), an approach that would yield an exponent of order $1/\sqrt{k}$ (using the version of the determinant

---

Date: November 16, 2009.

2000 *Mathematics Subject Classification.* Primary 11D85; Secondary 14G05.

Key words and phrases. Sum of k-th powers; Determinant method; diagonal form; integral points.

method developed in [6] and [7]), he studies rational points near the projective curve given by $x_1^k \pm x_2^k \pm x_3^k = 0$.

Our aim in this paper is to study the corresponding problem in four variables, using Heath-Brown's approach. The method works for arbitrary non-singular diagonal forms, so we state our main result in that generality. Let (a_1, a_2, a_3, a_4) be a quadruple of non-zero integers, $k \geq 3$ an integer, and N a positive integer. Let $\mathcal{R}(N, B)$ be the number of quadruples $(x_1, x_2, x_3, x_4) \in \mathbb{Z}^4$ satisfying

$$a_1 x_1^k + a_2 x_2^k + a_3 x_3^k + a_4 x_4^k = N \tag{3}$$

and $\max |x_i| \leq B$. We note that the trivial estimate $\mathcal{R}(N, B) = O(B^{2+\varepsilon})$ can be deduced easily using known results for Thue equations (see Proposition 5.1 below).

We call a solution $\mathbf{x}$ to (3) *special* if either $a_i x_i^k = N$ for some index i or $a_i x_i^k + a_j x_j^k = N$ for some pair of indices i, j. If $X \subset \mathbb{A}^4$ denotes the hypersurface defined by (3), then the special solutions are all contained in a proper subvariety of X, namely the union of all lines on X. Let $\mathcal{R}_0(N, B)$ be the number of non-special solutions to (3) satisfying $\max |x_i| \leq B$. We shall prove the following estimate.

Theorem 1.1. *For any $\varepsilon > 0$ we have*

$$\mathcal{R}_0(N, B) \ll_{N,\varepsilon} B^{16/(3\sqrt{3k})+\varepsilon}(B^{2/\sqrt{k}} + B^{1/\sqrt{k}+6/(k+3)}). \tag{4}$$

In particular, $\mathcal{R}(N, B) \ll_\varepsilon B^{1+\varepsilon}$ (that is, the special solutions dominate) for $k \geq 27$.

Remark 1.1. The exponent $16/(3\sqrt{3k})$ in Theorem 1.1 is to be compared with the exponent $3/k^{1/3}$ that could be obtained by applying the 'ordinary' determinant method of Heath-Brown [7, Thm. 15] in this case. The bound (4) is non-trivial for $k \geq 8$.

The estimate in Theorem 1.1 is proven by combining the ideas from [8] with recent results by Salberger [12] on integral points on affine surfaces.

In Sections 2 and 3 we adapt Heath-Brown's arguments to the four-variable case. As with other instances of the determinant method, the output is a number of auxiliary forms, allowing us to estimate $\mathcal{R}(N, B)$ through counting integral points of bounded height on a number of affine algebraic surfaces. In doing this, we use results by Salberger, discussed in Section 4, concerning the geometry of Fermat hypersurfaces. The proof of Theorem 1.1 is finished in Section 5.

It is implicit in Theorem 1.1 that N is fixed and small. If N is allowed to grow as $B \to \infty$, we have the following more precise estimate.

Theorem 1.2. *Suppose that $N = O(B^{\mu k})$, where $0 \leq \mu < k$. Then we have*

$$\mathcal{R}_0(N, B) \ll_\varepsilon B^{\frac{16}{3\sqrt{3k}}+\varepsilon} N^{\frac{24}{(3k(1-\mu))^{3/2}} - \frac{8}{(3k)^{3/2}}} \left(B^{2/\sqrt{k}} + B^{1/\sqrt{k}+6/(k+3)}\right) \tag{5}$$

for any $\varepsilon > 0$.

Note that, as in [8], the determinant method discussed in Sections 2 and 3 applies to any non-singular form. It is only in the later steps of the proof of Theorem 1.1 that we specialize to the case of a diagonal form.

The second result of the paper concerns the number $R_k(N)$ of representations of a positive integer N as a sum of four k-th powers

$$x_1^k + x_2^k + x_3^k + x_4^k = N, \tag{6}$$

where x_i are *non-negative* integers and $k \geq 3$. One easily proves, for example using Proposition 5.1 below, that $R_k(N) = O_\varepsilon(N^{2/k+\varepsilon})$. Hooley [9] has studied sums of four cubes, and proved the remarkable estimate $R_3(N) = O_\varepsilon(N^{11/18+\varepsilon})$. Wisdom [14, 15] extended Hooley's methods to prove that $R_k(N) = O_\varepsilon(N^{11/(6k)+\varepsilon})$ for odd integers $k \geq 3$. Our result is the following:

Theorem 1.3.
$$R_k(N) \ll_\varepsilon N^{1/k+2/k^{3/2}+\varepsilon}$$
for any $\varepsilon > 0$.

This estimate is non-trivial for $k \geq 5$, and sharper than Wisdom's for $k > 5$. Theorem 1.3, which follows rather promptly from Salberger's work [12], is proven in Section 6. In fact, with no extra work the method of proof yields the following more general theorem, to which Theorem 1.3 is a corollary.

Theorem 1.4. *Let* $k, \ell \geq 3$. *Let* $R_{k,\ell}(N)$ *be the number of solutions to the equation*

$$x_1^k + x_2^k + x_3^k + x_4^\ell = N, \tag{7}$$

in non-negative integers x_i. *Then we have the estimate*

$$R_{k,\ell}(N) \ll_\varepsilon N^{1/\ell+2/k^{3/2}+\varepsilon}$$

for any $\varepsilon > 0$.

The corresponding trivial estimate is $N^{1/\ell+1/k+\varepsilon}$. We also note that Wisdom [16] has proved that

$$R_{3,4}(N) = O_\varepsilon(N^{5/9+\varepsilon}) \text{ and } R_{3,5}(N) = O_\varepsilon(N^{47/90+\varepsilon}),$$

bounds which are sharper than the ones given by Theorem 1.4.

Notation. If $U \subseteq \mathbb{A}^n$ is a locally closed subset, let $U(\mathbb{Z})$ be the set of integral points in U. Then we define

$$U(\mathbb{Z}, B) = U(\mathbb{Z}) \cap [-B, B]^n$$

for any positive real number B, and

$$\mathcal{N}(U, B) = \#U(\mathbb{Z}, B).$$

We shall also use the notation

$$\mathcal{N}_+(U,B) = \#(U(\mathbb{Z}) \cap [0,B]^n).$$

Finally, we adopt the following convention for the O- and $\ll$-notation. The implied constants are allowed to depend upon the coefficients of the polynomial F (that is, on the a_i, in the case of a diagonal form) unless we indicate uniformity through the subscript k.

2. Parameterization of points near projective surfaces

In this section we generalize, in a completely straightforward fashion, some preparatory results in Heath-Brown's paper [8].

The proof of Lemma 1 in [8] generalizes readily to $\mathbb{R}^4$, to yield our first lemma.

Lemma 2.1. *Let $F(x_1, x_2, x_3, x_4)$ be a non-singular homogeneous polynomial of degree k. There is a natural number M_0, depending only on F, with the following property: if the unit cube $[-1,1]^3$ is partitioned into smaller cubes*

$$[a_1, a_1 + (M_0M)^{-1}] \times [a_2, a_2 + (M_0M)^{-1}] \times [a_3, a_3 + (M_0M)^{-1}],$$

for some positive integer M, then the number of such cubes containing a solution $(t_1, t_2, t_3) \in \mathbb{R}^3$ to the inequality

$$|F(t_1, t_2, t_3, 1)| \leq \frac{1}{M_0M} \tag{8}$$

is at most $O(M^2)$. Moreover, if S is such a cube containing a solution to (8), then for some index i we have

$$\left|\frac{\partial F}{\partial x_i}\right| \gg 1,$$

throughout S.

Following Heath-Brown, let us call such a cube S a 'good' cube. Let us also call a solution $\mathbf{t}$ to (8) a 'good' point.

For a good cube, we can prove the following result. Again, the proof is an easy generalization of that of [8, Lemma 2].

Lemma 2.2. *Retaining the notation of the previous lemma, suppose that*

$$S = [a_1, a_1 + (M_0M)^{-1}] \times [a_2, a_2 + (M_0M)^{-1}] \times [a_3, a_3 + (M_0M)^{-1}]$$

is a good cube. Suppose that $\left|\frac{\partial F}{\partial x_3}\right| \gg 1$.

If $(t_1, t_2, t_3) = (a_1 + u_1, a_2 + u_2, a_3 + u_3) \in S$, put

$$w = F(t_1, t_2, t_3, 1) - F(a_1, a_2, a_3, 1).$$

Then there exist, for each $m \in \mathbb{N}$, polynomials $\Phi_m(u_1, u_2, w)$ and $\Psi_m(u_1, u_2, u_3, w)$, such that Φ_m has no constant term and Ψ_m has no term of degree less than m, and such that the relation

$$u_3 = \Phi_m(u_1, u_2, w) + u_3\Psi_m(u_1, u_2, u_3, w)$$

holds throughout S. Moreover, Φ_m and Ψ_m have degree $O_m(1)$ and coefficients of size $O_m(1)$.

In other words, the lemma states that the relation

$$F(a_1 + u_1, a_2 + u_2, a_3 + u_3) = F(a_1, a_2, a_3) + w$$

defines u_3, approximately, as a function of u_1, u_2 and w. It may thus be viewed as a form of the Implicit function theorem.

3. Application of the determinant method

Let $F \in \mathbb{Z}[x_1, x_2, x_3, x_4]$ be a non-singular form of degree $k \geq 3$, N a positive integer, and $B \geq 1$ a real number. Our aim in this section is to exhibit a set $\mathcal{C}$ of homogeneous polynomials $A_i \in \mathbb{Z}[x_1, x_2, x_3, x_4]$ of the same degree δ, such that every solution $\mathbf{x} \in \mathbb{Z}^4 \cap [-B, B]^4$ to the inequality

$$|F(x_1, x_2, x_3, x_4)| \leq N \tag{9}$$

satisfies at least one of the equations $A_i(\mathbf{x}) = 0$. These polynomials shall be called auxiliary forms. Let $\mathcal{A}(F, N, B, \delta)$ be the smallest possible cardinality of such a collection $\mathcal{C}$ of auxiliary forms. This is a well-defined quantity since there are only finitely many solutions to (9). Our arguments will conclude in two different estimates for $\mathcal{A}(F, N, B, \delta)$. However, we begin with some considerations that apply to both situations.

Since we are only interested in the order of growth of $\mathcal{A}(F, N, B, \delta)$ as a function of B, it is clearly enough to consider solutions to (9) for which

$$|x_4| \geq \max(|x_1|, |x_2|, |x_3|).$$

We may even restrict ourselves to solutions satisfying

$$B/2 < \max_i(|x_i|) = |x_4| \leq B, \tag{10}$$

deducing the final estimate from this case by dyadic summation.

If we assume that

$$M \leq (B/2)^k M_0^{-1} N^{-1}, \tag{11}$$

then every solution $\mathbf{x}$ to (9) and (10) produces a good point

$$\mathbf{t} = (x_1/x_4, x_2/x_4, x_3/x_4),$$

which by Lemma 2.1 lies in some good cube. Thus, let

$$S = [a_1, a_1 + (M_0M)^{-1}] \times [a_2, a_2 + (M_0M)^{-1}] \times [a_3, a_3 + (M_0M)^{-1}]$$

be a good cube, and let $R = \{\mathbf{x}^{(j)}, 1 \le j \le J\}$ be the set of $\mathbf{x} \in \mathbb{Z}^4$ satisfying (9) and (10), and such that $\mathbf{t} \in S$. For now, let δ be any positive integer, and let $s = \binom{\delta+3}{3}$ be the number of different monomials in x_1, x_2, x_3, x_4 of degree δ. Consider the $s \times J$-matrix

$$\mathcal{A} = \left(f_i(\mathbf{x}^{(j)})\right),$$

where f_i runs over all monomials of degree δ. We shall prove that it is possible to choose M in such a way that $\operatorname{rank} \mathcal{A} < s$. This implies the existence of a homogeneous polynomial $A(x_1, x_2, x_3, x_4)$ of degree δ vanishing at all the $\mathbf{x}^{(j)}$.

If $J < s$, we are done. Otherwise, we proceed by choosing a subset of R of cardinality s, without loss of generality $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\}$, and evaluating the corresponding $s \times s$-subdeterminant

$$\Delta_1 = \det \left(f_i(\mathbf{x}^{(j)})\right)_{1 \le i,j \le s}.$$

Our aim is to prove that $|\Delta_1| < 1$. In that case, being an integer, Δ_1 has to vanish.

We have

$$\Delta_1 = \prod_{j=1}^{s} (x_4^{(j)})^{\delta} \Delta_2 \ll B^{s\delta} |\Delta_2|, \tag{12}$$

where $\Delta_2 = \det \left(f_i(t_1^{(j)}, t_2^{(j)}, t_3^{(j)}, 1)\right)$.

At this point, we make the 'variable change' suggested by Lemma 2.2. Suppose, without loss of generality, that $\partial F / \partial x_3 \gg 1$ throughout S. For $(t_1, t_2, t_3) \in S$, let u_1, u_2, u_3, w be as in Lemma 2.2. Furthermore, let $\xi = F(t_1, t_2, t_3, 1) = w + F(a_1, a_2, a_3, 1)$. Then, if $\mathbf{t}$ is a good point and $m \in \mathbb{N}$, we have

$$f_i(t_1, t_2, t_3, 1) = g_i(u_1, u_2, \xi) + O_m(M^{-1}(M^{-m} + (NB^{-k})^m)),$$

for some polynomials g_i of degree $O_m(1)$, since $|u_i| \le (M_0 M)^{-1}$ and $|\xi| \le N(B/2)^{-k}$. Here, g_i depends on the chosen cube S, but the size of the coefficients of g_i is bounded in terms of m. The above formula yields

$$\Delta_2 = \Delta_3 + O_m(M^{-1}(M^{-m} + (NB^{-k})^m)), \tag{13}$$

where $\Delta_3 = \det \left(g_i(u_1^{(j)}, u_2^{(j)}, \xi^{(j)})\right)$.

To estimate the determinant Δ_3, we shall use a variant of the argument in [1] where we take into account the fact that one of the variables, ξ, takes only small values. Let us recall the notation from [8]: let n, D, H be positive integers. Given real numbers $0 \le X_1, \dots, X_n \le 1$, we define the size $\|m_i\|$ of a monomial $m_i(x_1, \dots, x_n) = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ by

$$\|m_i\| = X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Furthermore, we enumerate the monomials $m_1, m_2, \dots$ in $x_1, \dots, x_n$ in such a way that $\|m_1\| \ge \|m_2\| \ge \cdots$. Finally, by abuse of notation, by the height $\|f\|$ of a polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ we mean the maximum modulus of its coefficients. Heath-Brown proves the following result.

Lemma 3.1. [8, Lemma 3] *Let $f_1, \ldots, f_H \in \mathbb{C}[x_1, \ldots, x_n]$ be polynomials of degree at most D. Let $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(H)} \in \mathbb{C}^n$ satisfy $|x_i^{(j)}| \leq X_i$ for all i and j. Then we have the estimate*

$$\det(f_i(\mathbf{x}^{(j)}))_{1\leq i,j\leq H} \ll_{H,D} (\max_i \|f_j\|)^H \prod_{j=1}^{H} \|m_i\|.$$

In the application of Lemma 3.1, we take $X_1 = X_2 = (M_0M)^{-1}$ and $X_3 = N(B/2)^{-k}$, according to our a priori bounds for $|u_1|$, $|u_2|$ and $|\xi|$, respectively.

Let the monomials $m_i(u_1, u_2, \xi)$ be defined as above. The strategy of our method is to ensure that for small i, m_i does not contain a positive power of ξ. In that way our determinant will behave almost as if we were considering points on a projective surface instead of points on an affine threefold. Thus, suppose that

$$M^\alpha = N^{-1}(B/2)^k, \tag{14}$$

where α is to be chosen properly. Then Lemma 3.1 yields an estimate

$$\Delta_3 \ll_{\delta,m} \prod_{i=1}^{s} \|m_i\| = M^{-f},$$

where f is an integer depending on α. To estimate f, we need to determine the first s monomials $m_1, \ldots, m_s$. Thus, let ν be the unique integer satisfying

$$M^{-(\nu-1)} < \|m_s\| \leq M^{-\nu}.$$

To determine the relationship between δ and ν, we note that

$$\sum_{\substack{n_1,n_2,n_3\geq 0\\ n_1+n_2+\alpha n_3\leq \nu-1}} 1 < s \leq \sum_{\substack{n_1,n_2,n_3\geq 0\\ n_1+n_2+\alpha n_3\leq \nu}} 1. \tag{15}$$

The left sum in (15), i.e. the number of integer points inside the tetrahedron $T_1 \subset \mathbb{R}^3$ defined by the inequalities

$$x \geq 0,\ y \geq 0,\ z \geq 0,\ x + y + \alpha z \leq \nu - 1,$$

can be interpreted as the volume of the three-dimensional body

$$S_1 = \bigcup_{\substack{n_1,n_2,n_3\geq 0\\ n_1+n_2+\alpha n_3\leq \nu-1}} [n_1, n_1 + 1] \times [n_2, n_2 + 1] \times [n_3, n_3 + 1].$$

Since $T_1 \subset S_1 \subset T_2$, where T_2 is the tetrahedron

$$x \geq 0,\ y \geq 0,\ z \geq 0,\ x - 1 + y - 1 + \alpha(z - 1) \leq \nu - 1,$$

we get the estimate

$$\frac{(\nu-1)^3}{6\alpha} = \mathrm{vol}(T_1) < \#(T_1 \cap \mathbb{Z}^3) < \mathrm{vol}(T_2) = \frac{(\nu+1+\alpha)^3}{6\alpha}. \tag{16}$$

Similarly, the right sum in (15) is the number of integer points inside the tetrahedron T_3 defined by the inequalities

$$x \geq 0,\ y \geq 0,\ z \geq 0,\ x - 1 + y - 1 + \alpha(z-1) \leq \nu,$$

so that,

$$\sum_{\substack{n_1,n_2,n_3 \geq 0 \\ n_1+n_2+\alpha n_3 \leq \nu}} 1 \leq \mathrm{vol}(T_3) = \frac{(\nu+2+\alpha)^3}{6\alpha}.$$

By the definition of s we conclude that

$$\delta = \frac{\nu}{\alpha^{1/3}} + O_\alpha(\nu^{2/3}). \tag{17}$$

A lower bound for f is given by the sum

$$\tilde{f} = \sum_{(n_1,n_2,n_3) \in T_1 \cap \mathbb{Z}^3} (n_1 + n_2 + \alpha n_3).$$

We can estimate $\tilde{f}$ from below by considering the integral

$$I = \int_{T_1} (x + y + \alpha z)\, dx\, dy\, dz = \frac{(\nu-1)^4}{8\alpha}$$

Since $T_1 \subset S_1$ and the integrand is increasing in x, y and z, we have

$$\begin{aligned} I &< \sum_{(n_1,n_2,n_3) \in T_1 \cap \mathbb{Z}^3} (n_1 + 1 + n_2 + 1 + \alpha(n_3 + 1)) \\ &= \tilde{f} + (2+\alpha)\#(T_1 \cap \mathbb{Z}^3). \end{aligned}$$

By (16) we conclude that

$$f > \frac{(\nu-1)^4}{8\alpha} - \frac{(2+\alpha)(\nu+1+\alpha)^3}{6\alpha} = \frac{\nu^4}{8\alpha} + O_\alpha(\nu^3). \tag{18}$$

Tracing our steps back to the estimates (12) and (13), and choosing $m = f$, we get the estimate

$$\Delta_1 \ll_\delta B^{s\delta} M^{-f} = B^\beta,$$

where, upon recalling the relation (14), we have

$$\beta = s\delta - \frac{f}{\alpha}\left(k - \frac{\log N}{\log B}\right).$$

Using (15), (17) and (18) this implies that

$$\beta = \left(\frac{1}{6\alpha^{4/3}} - \left(k - \frac{\log N}{\log B}\right)\frac{1}{8\alpha^2}\right)\nu^4 + O_\alpha(\nu^3). \tag{19}$$

Let us consider the simplest case, where N remains fixed as $B \to \infty$. Put

$$\alpha = (1-\lambda)\left(\frac{3k}{4}\right)^{3/2}$$

for some small real number $\lambda > 0$, so that the leading coefficient in (19) becomes negative. We may clearly assume that $\alpha > 1$, so that the requirement (11) is fulfilled for B larger than some constant depending only on F and N. It follows that there are constants δ_0 and B_0, depending only upon F, N and λ, such that $|\Delta_1| < 1$ provided that $B \geq B_0$ and $\delta \geq \delta_0$.

In this situation, as already explained, we obtain an auxiliary form of degree δ for each good cube S. By Lemma 2.1, the total number of good cubes is $O(M^2) = O(B^{2k/\alpha})$, and so this constitutes an upper estimate for $\mathcal{A}(F, N, B, \delta)$.

In particular, let ε be an arbitrary positive real number. Then, provided $\lambda < \lambda_0$, where λ_0 is some constant depending only upon F and ε, we have

$$\frac{2k}{\alpha} \leq \frac{16}{3\sqrt{3k}} + \varepsilon.$$

Thus we can summarize our findings so far in the following result

Proposition 3.1. *Let $F \in \mathbb{Z}[x_1, x_2, x_3, x_4]$ be a non-singular homogeneous polynomial of degree $k \geq 3$. Let $N \in \mathbb{N}$ be given. Then for any $\varepsilon > 0$, there is an integer δ, depending only on F and ε, such that*

$$\mathcal{A}(F, N, B, \delta) = O_{N,\varepsilon}(B^{16/(3\sqrt{3k})+\varepsilon}).$$

Next, let us allow N to vary, requiring merely that $N = O(B^{\mu k})$ for some constant $\mu < 1$. Let $\gamma = \log N / \log B$ and put

$$\alpha = (1-\lambda)\left(\frac{3(k-\gamma)}{4}\right)^{3/2}. \tag{20}$$

where $\lambda > 0$. Then (11) holds provided $\lambda \ll_\mu 1$ and $B \gg_\mu 1$. Again, the leading coefficient in (19) is negative, so for δ large enough we have

$$\mathcal{A}(F, N, B, \delta) \ll M^2 \approx B^{2k/\alpha} N^{-1/\alpha}.$$

From (20) we derive the estimate

$$\begin{aligned}
\frac{1}{\alpha} &= \left(\frac{4}{3k}\right)^{3/2} (1-\lambda)^{-1} \left(1 - \frac{\gamma}{k}\right)^{-3/2} \\
&= \left(\frac{4}{3k}\right)^{3/2} \left(1 + \lambda(1-\lambda)^{-1}\right) \left(1 + \frac{\gamma}{k}\left(1 - \frac{\gamma}{k}\right)^{-1}\right)^{3/2} \\
&< \left(\frac{4}{3k}\right)^{3/2} (1 + 2\lambda) \left(1 + \frac{3}{2(1-\mu)^{3/2}} \frac{\gamma}{k}\right)
\end{aligned}$$

In particular, for any $\varepsilon > 0$ we have

$$\frac{8}{3^{3/2}k^{3/2}} < \frac{1}{\alpha} < \frac{8}{3^{3/2}k^{3/2}} + \varepsilon + \gamma\frac{4}{3^{1/2}(1-\mu)^{3/2}k^{5/2}},$$

upon choosing $\lambda \ll_\varepsilon 1$. Using this estimate we have that

$$\mathcal{A}(F, N, B, \delta) \ll e^b,$$

where

$$\begin{aligned} b &= \frac{2k}{\alpha}\log B - \frac{1}{\alpha}\log N \\ &< \left(\frac{16}{3\sqrt{3k}} + \varepsilon\right)\log B + \left(\frac{24}{(3k(1-\mu))^{3/2}} - \frac{8}{(3k)^{3/2}}\right)\log N. \end{aligned}$$

Thus we have proven

Proposition 3.2. *Let $F \in \mathbb{Z}[x_1, x_2, x_3, x_4]$ be a non-singular homogeneous polynomial of degree $k \geq 3$. Let $N \in \mathbb{N}$ be given such that $N = O(B^{\mu k})$, where $\mu < 1$. Then for any $\varepsilon > 0$, there is an integer δ, depending only on F and ε, such that*

$$\mathcal{A}(F, N, B, \delta) = O_{\mu,\varepsilon}\left(B^{\frac{16}{3\sqrt{3k}}+\varepsilon} N^{\frac{24}{(3k(1-\mu))^{3/2}} - \frac{8}{(3k)^{3/2}}}\right).$$

4. Curves of low degree on Fermat threefolds

The following result was proven by Salberger [12, Thm. 8.1]:

Theorem 4.1. *Let K be an algebraically closed field of characteristic 0, $(a_0, \ldots, a_n)$ an $(n+1)$-tuple of non-zero elements of K and $X \subset \mathbb{P}^n_K$ the Fermat hypersurface given by $a_0x_0^k + \cdots + a_nx_n^k = 0$. Let $C \subset X$ be an irreducible closed curve of degree e that does not lie on any other hypersurface defined by a diagonal form $b_0x_0^k + \cdots b_nx_n^k$. Then the following holds:*

$$(n+1)(k-(n-1)) \leq nd + \frac{n(n-1)(e-3)}{2}.$$

In our special case of interest, we draw the following conclusion (cf. [12, Thm. 8.4]).

Proposition 4.1. *Let K be an algebraically closed field of characteristic 0, $(a_0, \ldots, a_4)$ a quintuple of non-zero elements of K and $X \subset \mathbb{P}^4_K$ the Fermat hypersurface given by $a_0x_0^k + \cdots + a_4x_4^k = 0$, where $k \geq 4$. Let $C \subset X$ be an irreducible closed curve. If*

$$\deg C < (k+3)/6,$$

then C is a line, and there is a partition

$$\{0, 1, 2, 3, 4\} = \{i_0, i_1\} \cup \{i_2, i_3, i_4\}$$

such that

$$a_{i_0}x_{i_0}^k + a_{i_1}x_{i_1}^k = a_{i_2}x_{i_2}^k + a_{i_3}x_{i_3}^k + a_{i_4}x_{i_4}^k = 0$$

on C.

Remark. These lines are called standard lines. The proposition extends (as soon as $k \geq 10$) the well-known fact that all lines contained in X are standard [3, Ex. 2.5.3].

Proof. We shall first prove the following statement (I): there exists a three-element subset $\{i_0, i_1, i_2\}$ of $\{0, 1, 2, 3, 4\}$ and a diagonal form $c_0 x_{i_0}^k + c_1 x_{i_1}^k + c_2 x_{i_2}^k$ that vanishes on C, with all $c_i \neq 0$.

By Theorem 4.1 there is a diagonal form $b_0 x_0^k + \cdots b_4 x_4^k$, linearly independent from $a_0 x_0^k + \cdots + a_4 x_4^k$, that vanishes on C. Choosing a suitable linear combination of the two forms, we can assume that either one or two of the coefficients b_i vanish. If there are exactly three non-zero coefficients we are done, so let us assume that there are four. By permuting the variables, we assume for the sake of simplicity that $b_4 = 0$ and $b_i \neq 0$ for $i < 4$.

Next, let $\pi : \mathbb{P}^4 \dashrightarrow \mathbb{P}^3$ be the rational map given by projection onto the first four coordinates. Let $Y \subset \mathbb{P}^3$ be the hypersurface given by $b_0 x_0^k + \cdots b_3 x_3^k = 0$. Then the image $\pi(C)$ is an irreducible curve $C' \subset Y$. Indeed, the image is either a point or a curve, but the first alternative would imply that C were a line containing the point $(0 : 0 : 0 : 0 : 1)$, which would contradict the fact that $a_4 \neq 0$. Furthermore we have $\deg C' \leq (k+3)/6 \leq (k+1)/3$, so by [12, Thm. 8.4], C' is a standard line. In other words, there is a partition $\{0, 1, 2, 3\} = \{j_0, j_1\} \cup \{j_2, j_3\}$ such that $b_{j_0} x_{j_0}^k + b_{j_1} x_{j_1}^k = b_{j_2} x_{j_2}^k + b_{j_3} x_{j_3}^k = 0$ on C. Choosing a suitable linear combination of the forms $a_0 x_0^k + \cdots + a_4 x_4^k$, $b_{j_0} x_{j_0}^k + b_{j_1} x_{j_1}^k$ and $b_{j_2} x_{j_2}^k + b_{j_3} x_{j_3}^k$, we get (I).

Having proven (I), we may assume, by permuting the variables, that the form $c_0 x_0^k + c_1 x_1^k + c_2 x_2^k$ vanishes on C. We shall prove that C is a line. To this end, let $\pi_1 : \mathbb{P}^4 \dashrightarrow \mathbb{P}^2$ be the projection onto the first three coordinates. Let $Z \subset \mathbb{P}^2$ be the subvariety given by $c_0 x_0^k + c_1 x_1^k + c_2 x_2^k = 0$. As above, $\pi_1(C)$ is either a point or an irreducible curve contained in Z. But this curve would have degree less than $(k+3)/6$, which would contradict the fact that Z is an irreducible curve of degree k. Therefore $\pi_1(C)$ is a single point, say $(y_0 : y_1 : 1)$ without loss of generality.

This means that C is contained in the plane $\Pi_1 \subset \mathbb{P}^4$ given by the equations $x_0 - y_0 x_2 = x_1 - y_1 x_2 = 0$. Inserting this into the equation for X, we infer that

$$a'_2 x_2^k + a_3 x_3^k + a_4 x_4^k = 0$$

on C, where $a'_2 = a_0 y_0^k + a_1 y_1^k + a_2$. If $a_2 = 0$, we infer that C is one of the k lines given by the equations

$$a_3 x_3^k + a_4 x_4^k = x_0 - y_0 x_2 = x_1 - y_1 x_2 = 0.$$

If $a_2 \neq 0$, then by the same argument as above, C is mapped to a point by the projection $\pi_2 : \mathbb{P}^4 \dashrightarrow \mathbb{P}^2$ onto the last three coordinates, which implies that C is contained in some plane $\Pi_2 \subset \mathbb{P}^4$, necessarily distinct from Π_1. C is then the line $\Pi_1 \cap \Pi_2$. It would now be easy to proceed by showing that C is one of the standard lines, but as remarked above, this is a known result. $\square$

5. Counting integral points on affine surfaces

From now on we consider the case of a diagonal form. Thus, let

$$F(x_1, x_2, x_3, x_4) = a_1 x_1^k + a_2 x_2^k + a_3 x_3^k + a_4 x_4^k,$$

where a_i are non-zero integers, let N be a positive integer and $B \geq 1$ a real number. Furthermore, let $X \subset \mathbb{A}^4$ be the hypersurface defined by $F(x_1, x_2, x_3, x_4) = N$. Let $V_i \subset \mathbb{A}^4$, for $1 \leq i \leq 4$, be the closed subvariety defined by

$$a_i x_i^k = N, \quad \sum_{j \neq i} a_j x_j^k = 0$$

and $W_{i,j}$, for $1 \leq i < j \leq 4$, be defined by

$$a_i x_i^k + a_j x_j^k = N, \quad \sum_{\ell \neq i,j} a_\ell x_\ell^k = 0,$$

The algebraic set

$$V = \left(\bigcup_{1 \leq i \leq 4} V_i \right) \cup \left(\bigcup_{1 \leq i < j \leq 4} W_{i,j} \right).$$

is precisely the union of all lines on X, by Proposition 4.1. The quantity we wish to estimate is then $\mathcal{R}_0(N, B) = \mathcal{N}(X_0, B)$, where $X_0 := X \setminus V$.

By Proposition 3.1 we know that every $\mathbf{x} \in X(\mathbb{Z}, B)$ satisfies

$$F(x_1, x_2, x_3, x_4) = N, \quad A_i(x_1, x_2, x_3, x_4) = 0, \tag{21}$$

for one of $O_{N,\varepsilon}(B^{16/(3\sqrt{3k})+\varepsilon})$ forms A_i of degree $O_\varepsilon(1)$.

Remark. Although Theorem 1.1, strictly speaking, is a corollary to Theorem 1.2, we shall only write out the proof of Theorem 1.1. If we would supply the more precise estimate of Proposition 3.2 at this point, we would obviously get a proof of Theorem 1.2.

Let $\tilde{Y} \subset \mathbb{A}^4$ be any one of the varieties defined by (21). Since A_i is homogeneous, it cannot vanish entirely on X, so the dimension of $\tilde{Y}$ is 2. Let Y be an irreducible component of $\tilde{Y}$. As we shall see shortly, we may assume that Y is in fact geometrically irreducible. Then, as $\bar{Y}$ is a closed subvariety of the non-singular hypersurface $\bar{X}$, where $\bar{X}, \bar{Y} \subset \mathbb{P}^4$ denote the respective projective closures, it follows from the Noether-Lefschetz theorem [5, pp. 180-1] that the degree d of $\bar{Y}$ is divisible by k.

It is then possible [10, Prop. 6.2] to find an affine projection $\pi : Y \to \mathbb{A}^3$ that is birational onto its image, and such that integral points of height at most B are mapped onto integral points of height at most

cB for some constant $c \ll_k 1$. Then $W = \pi(Y) \subset \mathbb{A}^3$ is an irreducible closed subvariety of dimension 2 and degree d, and $\pi^{-1}(\mathbf{x})$ consists of at most d points for any $\mathbf{x} \in W$.

Now we use the new version of the determinant method developed by Salberger. By [12, Cor. 7.3] we infer that there is a collection of $O_{k,\varepsilon}(B^{1/\sqrt{k}+\varepsilon})$ irreducible curves on W of degree $O_{k,\varepsilon}(1)$ such that all but $O_{k,\varepsilon}(B^{2/\sqrt{k}+\varepsilon})$ points of $W(\mathbb{Z}, B)$ lie on one of these curves. Pulling these curves and points back by π, we get $O_{k,\varepsilon}(B^{1/\sqrt{k}+\varepsilon})$ irreducible curves of bounded degree on Y, the union of which contains all but $O_{k,\varepsilon}(B^{2/\sqrt{k}+\varepsilon})$ points of $Y(\mathbb{Z}, B)$.

Concerning the case where Y is integral but not geometrically integral we can say more. Indeed, one can argue as in [13, Proof of Thm. 2.1] to conclude that all rational points on Y lie on a single curve, the sum of the degrees of the irreducible components of which is bounded in terms of k. Thus these irreducible components can be absorbed in the collection of curves and points of the previous paragraph.

To investigate the nature of such a curve, we shall use Proposition 4.1 on the hypersurface

$$\bar{X} = \{-Nx_0^k + a_1x_1^k + a_2x_2^k + a_3x_3^k + a_4x_4^k = 0\} \subset \mathbb{P}^4.$$

Any irreducible curve on X of degree less than $(k+3)/6$ gives rise to an irreducible curve of the same degree on $\bar{X}$, and must therefore in fact be one of the lines in V.

Since the number of irreducible components of a surface $\tilde{Y}$ as above is bounded in terms of k, we conclude that

$$X_0(\mathbb{Z}, B) \subseteq \left(\bigcup_C C(\mathbb{Z}, B)\right) \cup \left(\bigcup_{\mathbf{y}} \{\mathbf{y}\}\right), \tag{22}$$

where C runs over a collection of

$$O(B^{16/(3\sqrt{3k})+1/\sqrt{k}+\varepsilon})$$

irreducible curves of degree at least $(k+3)/6$, and $\mathbf{y}$ runs over a collection of

$$O(B^{16/(3\sqrt{3k})+2/\sqrt{k}+\varepsilon})$$

points.

To obtain the estimate (4), we now apply Pila's estimate [11]. If $C \subset \mathbb{A}^4$ is an irreducible curve of degree d, then we have

$$\mathcal{N}(C, B) \ll_{d,\varepsilon} B^{1/d+\varepsilon}. \tag{23}$$

Thus we conclude that

$$\mathcal{R}_0(N, B) \ll B^{16/(3\sqrt{3k})+1/\sqrt{k}+6/(k+3)+\varepsilon} + B^{16/(3\sqrt{3k})+2/\sqrt{k}+\varepsilon}. \tag{24}$$

To prove the last assertion in Theorem 1.1, we shall estimate $\mathcal{N}(V, B)$, using known bounds for Thue equations.

Proposition 5.1. *Let $a, b, h \in \mathbb{Z} \setminus \{0\}$, and let $k \geq 3$ be an integer. Then the number of integer solutions (x, y) to the equation $ax^k + by^k = h$ is $O(h^\varepsilon)$ for any $\varepsilon > 0$, where the implied constant depends only on k and ε.*

Proof. More precisely, the number of solutions is at most $C^{1+\omega(h)}$, where C is a constant depending only on k. This follows from Evertse's estimate [4, Cor. 2] for Thue-Mahler equations. Thus the proposition follows from the observation that $\omega(h) \ll \log h / \log\log h$. $\square$

We shall now estimate $\mathcal{N}(V_i, B)$ and $\mathcal{N}(W_{i,j}, B)$, where clearly it suffices to handle the case $(i, j) = (1, 2)$. Beginning with $\mathcal{N}(V_1, B)$, we have at most two choices for the value of x_1. For each fixed integer $x_4 \in [-B, B]$, there is then by Proposition 5.1 at most $O_\varepsilon(B^\varepsilon)$ integer solutions to the equation $a_2 x_2^k + a_3 x_3^k = -a_4 x_4^k$. We conclude that $\mathcal{N}(V_1, B) \ll_\varepsilon B^{1+\varepsilon}$.

Next we consider $\mathcal{N}(W_{1,2}, B)$. Here we have $O(B)$ choices for (x_3, x_4), and by Proposition 5.1 there are $O_\varepsilon(N^\varepsilon)$ possibilities for (x_1, x_2). Thus $\mathcal{N}(W_{1,2}, B) \ll_\varepsilon BN^\varepsilon$. In sum, we have

$$\mathcal{N}(V, B) \ll_\varepsilon (B(B^\varepsilon + N^\varepsilon)) \ll B^{1+\varepsilon}. \tag{25}$$

6. The sum of three k-th powers and an ℓ-th power

Let $X \subset \mathbb{A}^4$ be the hypersurface defined by the equation (7). We shall count integral points on hyperplane sections of X. Thus, for each integer $a \in [0, N^{1/\ell})$, let X_a be the intersection of X with the hyperplane given by $x_4 = a$. Viewed as a subvariety of $\mathbb{A}^3$, X_a is given by the equation

$$x_1^k + x_2^k + x_3^k = N - a^\ell.$$

Let $B = N^{1/k}$. It is then obvious that we have

$$R_{k,\ell}(N) \leq \sum_{0 \leq a < N^{1/\ell}} \mathcal{N}_+(X_a, B) + 1. \tag{26}$$

To estimate $\mathcal{N}_+(X_a, B)$, we use the results of Salberger [12]. By [12, Thm. 7.2] there is a collection $\mathcal{C}$ of (geometrically integral) curves $C \subset X_a$ of degree $O_{k,\varepsilon}(1)$, such that all but $O_{k,\varepsilon}(B^{2/\sqrt{k}+\varepsilon})$ points in $X_a(\mathbb{Z}, B)$ belong to one of the curves $C \in \mathcal{C}$, where $\#\mathcal{C} = O_{k,\varepsilon}(B^{1/\sqrt{k}+\varepsilon})$. In other words, we have

$$\mathcal{N}_+(X_a, B) \leq \sum_{C \in \mathcal{C}} \mathcal{N}_+(C, B) + O_{k,\varepsilon}(B^{2/\sqrt{k}+\varepsilon}). \tag{27}$$

Let $\bar{X}_a \subset \mathbb{P}^3$ be the projective closure of X_a, that is the Fermat hypersurface given by the equation

$$-(N - a^\ell)x_0^k + x_1^k + x_2^k + x_3^k = 0.$$

Since $\bar{X}_a$ is smooth, it follows from a theorem of Colliot-Thélène [2] that the number of geometrically integral curves on $\bar{X}_a$ that have degree at most $k-2$ is finite.

Using the results in [12], we can say more about the degrees of these curves. Indeed, by Theorem 8.4, any geometrically integral curve $\bar{X}_a$ that is not one of the standard lines has degree at least $(k+1)/3$.

This means the only curves $C \in \mathcal{C}$ that have degree less than $(k+1)/3$ are among the lines constituting the irreducible components of the subvarieties given by any the following pairs of equations:

$$x_1^k + x_2^k = x_3^k - (N - a^\ell) = 0, \tag{28}$$

$$x_1^k + x_3^k = x_2^k - (N - a^\ell) = 0, \tag{29}$$

$$x_2^k + x_3^k = x_1^k - (N - a^\ell) = 0. \tag{30}$$

However, each of the systems (28) - (30) has at most one solution in non-negative integers x_1, x_2, x_3, so the contribution to (27) from these lines is $O(1)$.

For the curves of higher degree, we use Pila's estimate [11]

$$\mathcal{N}(C, B) \ll_{\deg C, \varepsilon} B^{1/\deg C + \varepsilon}.$$

The bounded number of curves $C \in \mathcal{C}$ with $(k+1)/3 \leq \deg C \leq k-2$ thus contribute $O_{k,\varepsilon}(B^{3/(k+1)+\varepsilon})$ to (27), while the curves with degree at least $k-1$ contribute $O_{k,\varepsilon}(B^{1/\sqrt{k}+1/(k-1)+\varepsilon})$. In sum, we get

$$\begin{aligned} \mathcal{N}_+(X_a, B) &\ll_{k,\varepsilon} B^{2/\sqrt{k}+\varepsilon} + B^{1/\sqrt{k}+1/(k-1)+\varepsilon} + B^{3/(k+1)+\varepsilon} \\ &\ll_{k,\varepsilon} B^{2/\sqrt{k}+\varepsilon}, \end{aligned}$$

for $k \geq 3$. Inserting this into (26), we get

$$R_{k,\ell}(N) \ll_{k,\varepsilon} N^{1/\ell} B^{2/\sqrt{k}+\varepsilon} = N^{1/\ell + 2/k^{3/2} + \varepsilon'}.$$

This proves Theorem 1.4.

References

[1] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[2] J.-L. Colliot-Thélène. Appendix to: "The density of rational points on curves and surfaces" by D. R. Heath-Brown [Ann. of Math. (2) **155** (2002), no. 2, 553–595; MR1906595 (2003d:11091)]. *Ann. of Math. (2)*, 155(2):596–598, 2002.

[3] Olivier Debarre. *Higher-dimensional algebraic geometry*. Universitext. Springer-Verlag, New York, 2001.

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

[5] Robin Hartshorne. *Ample subvarieties of algebraic varieties*. Notes written in collaboration with C. Musili. Lecture Notes in Mathematics, Vol. 156. Springer-Verlag, Berlin, 1970.

[6] D. R. Heath-Brown. The density of rational points on curves and surfaces. *Ann. of Math. (2)*, 155(2):553–595, 2002.

[7] D. R. Heath-Brown. Counting rational points on algebraic varieties. In *Analytic number theory*, volume 1891 of *Lecture Notes in Math.*, pages 51–95. Springer, Berlin, 2006.

[8] D. R. Heath-Brown. Sums and differences of three kth powers. *J. Number Theory*, 129(6):1579–1594, 2009.

[9] C. Hooley. On the representations of a number as the sum of four cubes. I. *Proc. London Math. Soc. (3)*, 36(1):117–140, 1978.

[10] Oscar Marmon. The density of integral points on hypersurfaces of degree at least four. To appear, arXiv:0808.2910.

[11] J. Pila. Density of integral and rational points on varieties. *Astérisque*, (228):4, 183–187, 1995. Columbia University Number Theory Seminar (New York, 1992).

[12] Per Salberger. Counting rational points on projective varieties. Preprint, 2009.

[13] Per Salberger. Counting rational points on hypersurfaces of low dimension. *Ann. Sci. École Norm. Sup. (4)*, 38(1):93–115, 2005.

[14] Joel M. Wisdom. *On the representation of numbers as sums of powers.* PhD thesis, University of Michigan, 1998.

[15] Joel M. Wisdom. On the representations of a number as the sum of four fifth powers. *J. London Math. Soc. (2)*, 60(2):399–419, 1999.

[16] Joel M. Wisdom. On the representations of a number as the sum of three cubes and a fourth or fifth power. *Compositio Math.*, 121(1):55–78, 2000.

Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg

Mathematical Sciences, University of Gothenburg, SE-412 96 Gothenburg

E-mail address: marmon@chalmers.se