Equal Sums of Three Powers

T.D. Browning and D.R. Heath-Brown
Mathematical Institute, Oxford

1 Introduction

In this paper we shall be concerned with the number $N_d(B)$ of positive integer solutions to the equation

$$x_1^d + x_2^d + x_3^d = x_4^d + x_5^d + x_6^d, \tag{1.1}$$

in a region $\max x_i \le B$. Here d will be a fixed positive integer. There are $6B^3 + O(B^2)$ trivial solutions in which x_4, x_5, x_6 are a permutation of x_1, x_2, x_3. We write $N_d^{(0)}(B)$ for the number of non-trivial solutions, and our task is to estimate this quantity. When $d = 2$ it is a simple exercise with the circle method to show that $N_d(B) \sim cB^4$ for an appropriate positive constant c, so that the non-trivial solutions dominate the trivial ones. For larger values of d, it follows from Hua's inequality [11] that

$$N_d(B) \ll_\varepsilon B^{7/2+\varepsilon}$$

for any fixed $\varepsilon > 0$, and recent work of the second author [8; Theorem 13] improves this for $d \ge 24$ to

$$N_d(B) \ll_\varepsilon B^{3+\theta+\varepsilon}, \quad \theta = \frac{2}{\sqrt{d}} + \frac{2}{d-1}.$$

(Here, and throughout the paper, we shall permit any implied constants to depend on the degree d.) In the case $d = 3$ it has been shown by Hooley [9], and independently by Heath-Brown [7], that

$$N_3(B) \ll_\varepsilon B^{3+\varepsilon},$$

providing that the Hasse-Weil L-functions of certain cubic 3-folds obey the standard analytic continuation and Riemann Hypotheses. The above results give highly non-trivial bounds for $N_d(B)$, but none the less fail to show whether or not the diagonal solutions dominate the non-diagonal ones.

We can now state the principal result of this paper.

Theorem *For any $d \ge 25$ and any $\varepsilon > 0$ we have*

$$N_d^{(0)}(B) \ll_\varepsilon B^{3-1/24+\varepsilon} + B^{5/2+5/2\sqrt{d}+11/6(d-1)+\varepsilon} + B^{41/16+5/2\sqrt{d}}.$$

In particular we have

$$N_d^{(0)}(B) = o(B^3)$$

for $d \geq 33$.

The equation (1.1) defines a non-singular hypersurface $X_d \subset \mathbb{P}^5$, say, and it would be little trouble to prove the same upper bound for the number of non-trivial integral points $\mathbf{x} \in \mathbb{Z}^6$ lying in the intersection $X_d \cap [-B, B]^6$, where we now consider $\mathbf{x}$ to be 'trivial' if it lies on any plane contained in X_d. For $d \geq 7$, the hypersurface X_d is of general type, and we have a conjecture of Batyrev and Manin [1] which in this setting states that any integral points must lie on one of a finite number of proper subvarieties of X_d. Certainly we do not know of even a single non-trivial solution to (1.1), as soon as $d \geq 7$.

For small d, various lower bounds for $N_d^{(0)}(B)$ are known. Thus it follows from work of Hooley [10; Theorem 3] that $N_3^{(0)}(B) \gg B^3$. For $d = 4$ the identity

$$\begin{aligned}(23x + 37y)^4 &+ (40x - 6y)^4 + (17x - 63y)^4 \\ &= \quad (23x - 57y)^4 + (40x + 6y)^4 + (17x + 63y)^4\end{aligned}$$

which Dickson [3; Vol. II, p. 710] attributes to V.G. Tariste, shows that we have $N_4^{(0)}(B) \gg B^2$. In general, as soon as (1.1) has a non-trivial solution we may use multiples of it to show that $N_d^{(0)}(B) \gg B$. For $d = 5$ and $d = 6$ we have the examples

$$49^5 + 75^5 + 107^5 = 39^5 + 92^5 + 100^5$$

and

$$3^6 + 19^6 + 22^6 = 10^6 + 15^6 + 23^6.$$

Indeed when $d = 5$ or $d = 6$, parametric solutions are known, which provide infinitely many essentially different solutions. However these are not sufficiently dense that one can use them to improve on the lower bound $N_d^{(0)}(B) \gg B$.

Our problem can be phrased in various alternative ways. One may set

$$S(\alpha) = \sum_{n \leq B} e(n\alpha)$$

and ask about the behaviour of the sixth power mean $I_{d,6}(B) = N_d(B)$, where

$$I_{d,s}(B) = \int_0^1 |S(\alpha)|^s d\alpha.$$

Theorem 1 shows that $I_{d,6}(B)$ is asymptotically $6B^3$ for $d \geq 33$. There have been extensive investigations of moments of this type, for general s, for sums analogous to $S(\alpha)$ in which the variable n runs over smooth numbers (see Vaughan and Wooley [14], for example). This restriction allowed sharper upper bounds than could be obtained for $S(\alpha)$ itself. However we now see that, at least for large d, one can establish the optimal estimate for $s = 6$, and that it is not necessary to restrict attention to smooth numbers.

One can also interpret these results in terms of the function

$$r(n) = \#\{(x_1, x_2, x_3) \in \mathbb{N}^3 : n = x_1^d + x_2^d + x_3^d\}.$$

Our theorem has the following consequence.

Corollary *When $d \geq 33$ we have*

$$\sum_{n \leq x} r(n)^2 \sim 6cx^{3/d},$$

with

$$c = \Gamma(1 + \frac{1}{d})^3 / \Gamma(1 + \frac{3}{d}).$$

For $d \geq 33$ there are asymptotically $\frac{1}{6}cx^{3/d}$ integers $n \leq x$ which are sums of three d-th powers, and almost all of these have essentially just one representation.

The significance of the constant c is that

$$\sum_{n \leq x} r(n) \sim cx^{3/d} \tag{1.2}$$

for any $d \geq 1$. From this it follows immediately that the number of representable integers n is most $\{\frac{1}{6}c + o(1)\}x^{3/d}$. The problem of lower bounds on the number of such n is one which has attracted a lot of attention over many years, and hitherto the best result available for large d was that due to Wooley [16; Theorem 1.3]. This latter result shows that the number of integers $n \leq x$ representable as sums of three d-th powers is $\gg x^{3/d-\delta}$, with $\delta = e^{-d/17}$.

Notation. We shall follow common practice in allowing the small positive quantity ε to take different values at different points of the argument.

Acknowledgements. While working on this paper, the first author was supported by EPSRC Grant number GR/R93155/01. Parts of the paper were prepared while the second author was a guest of the Mathematics Department of the University of Hong Kong, and of the American Institute of Mathematics. The support of all these bodies is gratefully acknowledged.

2 Preliminaries

Our arguments will require various preliminary results which we gather here. In stating these, and elsewhere in the paper, we shall write $|\mathbf{x}|$ for the Euclidean length $(\sum_1^n x_i^2)^{1/2}$ of the vector $\mathbf{x}$. Since n will be assumed fixed, this norm is equivalent to that given by $\max |x_i|$. By $\mathbf{x} \ll 1$ we shall mean that $|\mathbf{x}| \ll 1$.

We shall require a number of facts from the geometry of numbers, concerning lattices. We shall be concerned with lattices $\Lambda \subseteq \mathbb{Z}^n \subset \mathbb{R}^n$. If Λ has dimension (or rank) r, we say that Λ is 'primitive' if it has a basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ which can be extended to a basis of $\mathbb{Z}^n$. If there is one basis which can be so extended, then any basis has such an extension. A necessary and sufficient condition for Λ to be primitive is that it should not be properly contained in any other r-dimensional sublattice of $\mathbb{Z}^n$.

If $\mathbf{b}_1, \ldots, \mathbf{b}_r$ is a basis for Λ we define $\det(\Lambda)$ to be the r-dimensional volume of the parallelepiped generated by $\mathbf{b}_1, \ldots, \mathbf{b}_r$. This is independent of the choice of the basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$. It follows immediately that

$$\det(\Lambda) \leq \prod_{i=1}^{r} |\mathbf{b}_i|. \tag{2.1}$$

If $\Lambda \subseteq \mathsf{M}$ are two lattices of the same dimension, then $\det(\Lambda) \geq \det(\mathsf{M})$.

We define the dual lattice Λ^* to be

$$\Lambda^* = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x}.\mathbf{y} = 0 \ \forall \mathbf{y} \in \Lambda\}.$$

Then Λ^* will always be primitive, with dimension $n - r$. Moreover if Λ is primitive, then $(\Lambda^*)^* = \Lambda$, and

$$\det(\Lambda^*) = \det(\Lambda).$$

All these facts are established in Heath-Brown [6; §2], the final assertion being [6; Lemma 1].

Henceforth, whenever we talk of a lattice, we shall mean a primitive lattice, unless we explicitly say otherwise. The facts we shall need later are gathered together in the following lemma.

Lemma 1

(i) For any primitive vector $\mathbf{c} \in \mathbb{Z}^n$ *the set* $\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{c}.\mathbf{x} = 0\}$ *is a lattice of dimension* $n-1$ *and determinant* $\det(\Lambda) = |\mathbf{c}|$.

(ii) Let $\Lambda \subseteq \mathbb{Z}^n$ *be a lattice of dimension* m. *Then* Λ *has a basis* $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(m)}$ *such that if one writes* $\mathbf{x} \in \Lambda$ *as* $\mathbf{x} = \sum_j \lambda_j \mathbf{b}^{(j)}$, *then*

$$\lambda_j \ll |\mathbf{x}|/|\mathbf{b}^{(j)}|. \tag{2.2}$$

Moreover one has $|\mathbf{b}^{(1)}| \leq \ldots \leq |\mathbf{b}^{(m)}|$ *and*

$$\det(\Lambda) \leq \prod_{j=1}^{m} |\mathbf{b}^{(j)}| \ll \det(\Lambda). \tag{2.3}$$

(iii) Let $\Lambda \subseteq \mathbb{Z}^n$ *be a lattice of dimension* m. *Then if* $R \gg 1$, *the sphere* $|\mathbf{x}| \leq R$ *contains* $O(R^m/\det(\Lambda) + R^{m-1})$ *points of* Λ.

(iv) If $\mathsf{\Lambda}$ and the basis $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(m)}$ are as in part (ii), then, for any integer $1 \le k \le m$, the lattice

$$\mathsf{M} = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x}.\mathbf{b}^{(1)} = \ldots = \mathbf{x}.\mathbf{b}^{(k)} = 0\}$$

has

$$\prod_{j=1}^{k} |\mathbf{b}^{(j)}| \ll \det(\mathsf{M}) \le \prod_{j=1}^{k} |\mathbf{b}^{(j)}|.$$

For the first assertion we merely note that $\mathsf{\Lambda} = \mathsf{M}^*$, where $\mathsf{M} = \langle \mathbf{c} \rangle$. Thus M is a lattice of dimension 1 and determinant $|\mathbf{c}|$.

For statement (ii) we note that Davenport [2; Lemma 5] shows the existence of a basis $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(m)}$ with the property (2.2), and in the course of the proof he constructs it in such a way that $|\mathbf{b}^{(1)}| \le \ldots \le |\mathbf{b}^{(m)}|$. Moreover, it is shown [2; (14)] that

$$\prod_{j=1}^{m} |\mathbf{b}^{(j)}| \ll \det(\mathsf{\Lambda}),$$

which gives (2.3), since the first inequality is merely a restatement of (2.1).

To prove statement (iii) we apply part (ii). This shows that the number of vectors to be counted is at most the number of m-tuples $(\lambda_1, \ldots, \lambda_m) \in \mathbb{Z}^m$ such that $\lambda_j \ll R/|\mathbf{b}^{(j)}|$ for every index j. Since $|\mathbf{b}^{(j)}| \ge 1$ this gives us a bound

$$\ll \prod_{j=1}^{m} (1 + \frac{R}{|\mathbf{b}^{(j)}|}) \ll R^{m-1} + \frac{R^m}{\prod_j |\mathbf{b}^{(j)}|} \ll R^{m-1} + \frac{R^m}{\det(\mathsf{\Lambda})},$$

as required.

Turning to statement (iv), we first observe that M is a primitive lattice of dimension $n-k$, so that (ii) provides a basis $\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(k)}$ for the dual lattice M^*, such that

$$\prod_{j=1}^{k} |\mathbf{c}^{(j)}| \asymp \det(\mathsf{M}^*) = \det(\mathsf{M}). \tag{2.4}$$

However it is clear that $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(k)} \in \mathsf{M}^*$, whence

$$\mathsf{L} = \langle \mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(k)} \rangle \subseteq \langle \mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(k)} \rangle = \mathsf{M}^*,$$

say. Since $\mathsf{\Lambda}$ was assumed to be primitive, it automatically follows that L is primitive, and so $\mathsf{L} = \mathsf{M}^*$. But then we have

$$\det(\mathsf{M}) = \det(\mathsf{M}^*) \le \prod_{j=1}^{k} |\mathbf{b}^{(j)}|$$

by (2.1), as required for the upper bound in part (iv). Arguing similarly, we may also conclude that

$$\mathsf{\Lambda} = \langle \mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(m)} \rangle = \langle \mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(k)}, \mathbf{b}^{(k+1)}, \ldots, \mathbf{b}^{(m)} \rangle,$$

since $\mathsf{L} = \mathsf{M}^*$. Thus

$$\prod_{j=1}^{m} |\mathbf{b}^{(j)}| \ll \det(\Lambda) \le \left(\prod_{j=1}^{k} |\mathbf{c}^{(j)}|\right) \left(\prod_{j=k+1}^{m} |\mathbf{b}^{(j)}|\right),$$

which yields

$$\prod_{j=1}^{k} |\mathbf{b}^{(j)}| \ll \prod_{j=1}^{k} |\mathbf{c}^{(j)}|.$$

We therefore obtain the desired lower bound in part (iv), via (2.4). This completes the proof of Lemma 1.

We shall also require various results concerning the number of points lying on algebraic curves and surfaces. We shall take these from the second author's recent work [8]. For a hypersurface V in $\mathbb{P}^{n-1}$, defined as the zero locus of an absolutely irreducible form $F(\mathbf{x}) \in \overline{\mathbb{Q}}[x_1, \ldots, x_n]$, we shall write

$$S(F; B_1, \ldots, B_n) = \\ \{\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} : F(\mathbf{x}) = 0,\ \text{h.c.f.}(x_1, \ldots, x_n) = 1,\ |x_i| \le B_i,\ (1 \le i \le n)\}.$$

When $B_1 = \ldots = B_n = B$ we merely write

$$S(F; B_1, \ldots, B_n) = S(F; B).$$

Furthermore, let $\|F\|$ denote the maximum modulus of the coefficients of F, let $V = B_1 \ldots B_n$, and let

$$T = \max\{B_1^{e_1} \ldots B_n^{e_n}\},$$

where the maximum is taken over all n-tuples $(e_1, \ldots, e_n)$ for which the monomial $x_1^{e_1} \ldots x_n^{e_n}$ occurs in $F(\mathbf{x})$ with non-zero coefficient. With this notation we have the following results.

Lemma 2 *Let $F(\mathbf{x}) \in \mathbb{Q}[x_1, \ldots, x_n]$ be a non-zero form of degree $D \ge 1$. Then*

$$\#S(F; B) \ll_{D,n} B^{n-1}.$$

Lemma 3 *Let $F(\mathbf{x}) \in \mathbb{Q}[x_1, \ldots, x_4]$ be an absolutely irreducible form of degree D, and let $\varepsilon > 0$ be given. Then every point in the set $S(F; B_1, \ldots, B_4)$ lies on one of at most*

$$\ll_{\varepsilon,D} (V||F||)^{\varepsilon} (V^D/T)^{D^{-3/2}}$$

curves of degree $O_D(1)$, contained in the surface $F = 0$.

Lemma 4 *Let $F(\mathbf{x}) \in \mathbb{Q}[x_1, \ldots, x_n]$ be an absolutely irreducible form of degree $D \ge 2$, and let $\varepsilon > 0$ be given. Then*

$$\#S(F; B) \ll_{\varepsilon,D,n} B^{n-5/3+\varepsilon}.$$

Lemma 5 *Let $C \subset \mathbb{P}^3$ be an irreducible curve of degree $D \geq 2$, not necessarily defined over the rationals, and assume that $B_1 \geq B_2 \geq B_3 \geq B_4 \geq 1$. Then C has*

$$\ll_{\varepsilon,D} (B_1B_3)^{1/D+\varepsilon}$$

points $\mathbf{x} \in \mathbb{Z}^4$ lying in the region $|x_i| \leq B_i$ for $1 \leq i \leq 4$.

We should emphasise that the implied constants in the above results are independent of F. This fact will be crucial in our arguments.

Lemma 2 is [8; Theorem 1], while Lemma 3 is merely a restatement of [8; Theorem 14] for the case $n = 4$. To establish Lemma 4 we note that Theorem B of Pila [13] yields a bound $O_{\varepsilon,D,n}(B^{n-2+1/D+\varepsilon})$ for forms of degree D, while [8; Theorem 2] yields the stronger bound $O_{\varepsilon,n}(B^{n-2+\varepsilon})$ when $D = 2$.

Lemma 5 is just a generalisation of [8; Theorem 5]. We therefore recall the proof of [8; Lemma 7], the thrust of which is concerned with finding certain integer vectors $\mathbf{y}$ and $\mathbf{c}$. These have lengths $|\mathbf{y}|, |\mathbf{c}| \ll 1$, and satisfy $\mathbf{y}.\mathbf{c} \neq 0$. Moreover the projection of C along $\mathbf{y}$ onto the plane $\mathbf{c}.\mathbf{x} = 0$ is an irreducible plane curve of degree D. An examination of the proof shows that the vector $\mathbf{y}$ can be chosen in such a way that all of the components are non-zero. Therefore the vector $\mathbf{c} = (0, 0, 0, 1)$ cannot be perpendicular to $\mathbf{y}$, and the projection π of C along $\mathbf{y}$ onto the plane $x_4 = 0$ produces an irreducible plane curve $\pi(C)$ of degree D. But then if $\mathbf{x} \in \mathbb{Z}^4$, with $|x_i| \leq B_i$, we have

$$(\mathbf{y}.\mathbf{c})\pi(\mathbf{x}) = y_4\mathbf{x} - x_4\mathbf{y} \in \mathbb{Z}^4$$

as in the proof of [8; Theorem 5], and it suffices to count primitive integer points $\mathbf{z}$ contained in the plane $z_4 = 0$, which lie on the curve $\pi(C)$ and in the region $|z_i| \leq |y_4x_i| + |y_ix_4| \ll B_i$ for $1 \leq i \leq 4$. Since $\pi(C)$ is irreducible, the form defining it must contain a monomial of the shape $z_1^a z_2^b$ with $a+b = D$. Therefore an application of [8; Theorem 3], with $T \geq B_2^D$, completes the proof of Lemma 5.

3 Singular Hyperplane Sections

It will become clear in Section 5 that we can get sharper results for non-singular surfaces than we can for singular ones. Our surfaces will arise from repeated hyperplane sections of the variety

$$F(\mathbf{x}) = x_1^d + x_2^d + x_3^d - (x_4^d + x_5^d + x_6^d) = 0,$$

and so it will be important to understand when a hyperplane section of this can be singular. In fact it will be more efficient to take hyperplane sections in only five variables. Thus we shall study the intersection of $F(\mathbf{x}) = 0$ with the hyperplane $\mathbf{a}.\mathbf{x} = 0$, where $\mathbf{a}$ is a primitive integer vector for which $a_6 = 0$. If a_1, say, is non-zero, we can eliminate x_1 from the two equations to produce a single equation in $x_2, \ldots, x_6$. We shall write $F_{\mathbf{a}}$ for the threefold produced by such an elimination process.

It will be convenient to make a change of variable, writing

$$x_i = y_i, \quad (1 \le i \le 3) \quad \text{and} \quad x_i = \omega y_i \quad (4 \le i \le 6),$$

where ω is a fixed d-th root of -1. This transforms $F(\mathbf{x})$ into

$$G(\mathbf{y}) = y_1^d + y_2^d + y_3^d + y_4^d + y_5^d + y_6^d.$$

Similarly we shall write

$$a_i = b_i, \quad (1 \le i \le 3) \quad \text{and} \quad a_i = \omega^{-1} b_i \quad (4 \le i \le 6), \tag{3.1}$$

so that $\mathbf{a}.\mathbf{x} = \mathbf{b}.\mathbf{y}$ Thus we now investigate vectors $\mathbf{b}$ for which $G_{\mathbf{b}}$ is singular.

If $G_{\mathbf{b}}$ is singular there is a vector $\mathbf{y}$ with $\nabla G(\mathbf{y}) = (d-1)\mathbf{b}$, so that $y_i^{d-1} = b_i$. Let us write $d-1 = h$ and $b_i^d = r_i$, for convenience. Note that, in view of the substitution (3.1), we will have $r_i \in \mathbb{Z}$. We have $r_i = y_i^{hd}$, so that we may fix an h-th root of r_i by setting $r_i^{1/h} = y_i^d$. We now define an equivalence relation on the non-zero r_i by saying that $r_i \sim r_j$ if and only if $(r_i/r_j)^{1/h} \in \mathbb{Q}(\varepsilon)$, where ε is a primitive h-th root of unity. Since the only primes which ramify in $\mathbb{Q}(\varepsilon)$ are those dividing h, the relation $r_i \sim r_j$ implies that $r_i = r_j q^h h_0$, where $q \in \mathbb{Q}$ and h_0 is one of the finite set of integers composed of prime factors of h, each taken with an exponent between 0 and $h-1$.

Since $F(\mathbf{y}) = 0$ and $a_6 = b_6 = 0$, we have

$$\sum_{i=1}^{5} r_i^{1/h} = 0.$$

We group the terms of this relation into equivalence classes, to produce an equation of the form

$$\sum \mu_i s_i^{1/h} = 0, \tag{3.2}$$

where the s_i are equivalence class representatives and $\mu_i \in \mathbb{Q}(\varepsilon)$. Terms with $r_i = 0$ are omitted.

We claim that the $s_i^{1/h}$ are linearly independent over $\mathbb{Q}(\varepsilon)$, so that $\mu_i = 0$ for each i. If not, choose a minimal linearly dependent set, $s_1^{1/h}, \ldots, s_r^{1/h}$, say, and use it to write

$$1 = \sum_{i=2}^{r} \nu_i t_i^{1/h},$$

where for each i we have

$$\nu_i \in \mathbb{Q}(\varepsilon), \quad \nu_i \ne 0, \quad t_i = s_i s_1^{-1}, \quad t_i^{1/h} \notin \mathbb{Q}(\varepsilon).$$

For any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}})$, write

$$\sigma(\nu_i t_i^{1/h}) = \nu_i' t_i^{1/h},$$

with $\nu_i' \in \mathbb{Q}(\varepsilon)$. Since $\nu_2 t_2^{1/h} \notin \mathbb{Q}$, there is a σ such that $\nu_2' \neq \nu_2$. By subtraction we then obtain a non-trivial relation

$$\sum_{i=2}^{r} (\nu_i' - \nu_i) t_i^{1/h} = 0,$$

which contradicts the supposed minimality of the set $s_1^{1/h}, \ldots, s_r^{1/h}$.

It therefore follows that $\mu_i = 0$ for each i in (3.2), so that

$$\sum_{i \in \mathcal{E}} r_i^{1/h} = 0 \tag{3.3}$$

for each equivalence class $\mathcal{E}$. In particular, we see that each equivalence class must have size at least 2, and that the equivalence classes form a partition of $\{2, 3, 4, 5, 6\}$. It follows that there are at most 2 equivalence classes. Write each element a_i in the form

$$a_i = e_i^h h_i f_i,$$

where h_i and f_i are h-th power free, $h_i | h^{h-1}$, and $(f_i, h) = 1$. We shall assume that $f_i > 0$, but not necessarily that e_i or h_i is positive. Then

$$r_i = b_i^d = \pm a_i^d = \pm a_i^{h+1} = \pm (e_i^{h+1} h_i f_i)^h h_i f_i,$$

and hence all elements r_i from the same equivalence class will have the same value for f_i, which we shall call f, say. The relation (3.3) now produces

$$\sum_{i \in \mathcal{E}} \lambda_i e_i^d = 0, \tag{3.4}$$

where $\lambda_i = h_i^{d/h} \varepsilon_i$, for some $2h$-th root of unity ε_i. Thus there are only $O_d(1)$ admissible sets of coefficients λ_i.

We proceed to investigate the number of possible vectors $\mathbf{a}$ in the region $|\mathbf{a}| \ll Y$, for which $a_6 = 0$. For each equivalence class we will have $e_i \ll (Y/f)^{1/h}$. Moreover if $\#\mathcal{E} = E$, say, then $E - 1$ of the variables determine the remaining one, by (3.4). Thus (3.4) has $O((Y/f)^{(E-1)/h})$ solutions. It follows that the number of possibilities for a_i, for $i \in \mathcal{E}$, is

$$\ll \sum_{f \ll Y} (Y/f)^{(E-1)/h} \ll Y$$

providing that $E - 1 < h$. We shall assume that $d \geq 6$, so that this latter condition holds automatically. Since there are at most 2 equivalence classes it follows that there are at most $O(Y^2)$ possible vectors $\mathbf{a}$ in total. We summarise our conclusion in the following lemma.

Lemma 6 *Suppose that $d \geq 6$. Then there are $O(Y^2)$ primitive vectors $\mathbf{a} \in \mathbb{Z}^6 \setminus \{0\}$ in the cube $\max |a_i| \leq Y$ for which $a_6 = 0$ and $F_{\mathbf{a}}$ is singular.*

4 Curves of Small Degree

In this section we shall investigate curves of degree at most three, lying in the hypersurface $F(\mathbf{x}) = 0$. We shall work over $\overline{\mathbb{Q}}$, so that we may consider the hypersurface defined by the form

$$G(\mathbf{y}) = y_1^d + y_2^d + y_3^d + y_4^d + y_5^d + y_6^d.$$

Our key tool is the following result.

Lemma 7 *Let $n \geq 2$ and let $f_1(z), \ldots, f_n(z)$ be meromorphic functions, which do not all vanish simultaneously. Suppose that $d \geq (n-1)^2$. Then if*

$$\sum_{j=1}^{n} f_j(z)^d = 0$$

holds identically, there must be two functions f_i, f_j which are proportional to each other.

This is due to Green [5; §7]. The proof uses Nevanlinna theory. However for the case in which the functions are polynomials one can give an elementary treatment following the method of Newman and Slater [12]. (We take this opportunity to record an apparent misprint in [12]. The proof of the theorem on page 481 of [12] seems to yield the condition $\deg R \geq n - (k^2 - k)$, rather than $\deg R \geq n - (k^2 - k)/2$).

Any curve of degree at most 3 is either a curve of genus zero, or is a plane cubic curve. In the first case the curve can be parameterized by polynomials in one variable. Since the polynomials can be assumed to have no common factor over $\overline{\mathbb{Q}}$ they will not vanish simultaneously. In the second case the curve can be parameterized using the Weierstrass elliptic function $\mathcal{P}(z)$. Thus for any such curve lying in the hypersurface $F = 0$ we can produce a set of functions

$$f_i(z) = A_i \mathcal{P}'(z) + B_i \mathcal{P}(z) + C_i, \quad (1 \leq i \leq 6),$$

not all proportional, satisfying the equation

$$\sum_{j=1}^{6} f_j(z)^d = 0$$

identically. Moreover, we must have

$$\mathrm{rank} \begin{pmatrix} A_1 & \ldots & A_6 \\ B_1 & \ldots & B_6 \\ C_1 & \ldots & C_6 \end{pmatrix} = 3,$$

since otherwise the vector $(f_1(z), \ldots, f_6(z))$ would run over a line, rather than a cubic curve. It follows that there exist $i < j < k$ for which the matrix

$$\begin{pmatrix} A_i & A_j & A_k \\ B_i & B_j & B_k \\ C_i & C_j & C_k \end{pmatrix}$$

is non-singular. Thus the simultaneous equations

$$\begin{aligned} f_i(z_0) &= A_i\mathcal{P}'(z_0)+B_i\mathcal{P}(z_0)+C_i \;=\; 0\\ f_j(z_0) &= A_j\mathcal{P}'(z_0)+B_j\mathcal{P}(z_0)+C_j \;=\; 0\\ f_k(z_0) &= A_k\mathcal{P}'(z_0)+B_k\mathcal{P}(z_0)+C_k \;=\; 0 \end{aligned}$$

have no solution. It follows that there is no $z_0\in\mathbb{C}$ at which $f_i(z), f_j(z), f_k(z)$ all vanish. We therefore deduce that any curve of degree at most 3 is parameterized by meromorphic functions which never vanish simultaneously.

We may now apply Lemma 7 to our situation. Suppose that $d\geq 25$. Then for any C of degree at most 3, lying in the hypersurface $F(\mathbf{x})=0$, we can produce a set of non-zero functions $f_1,\ldots,f_n$ with $2\leq n\leq 6$ such that the f_j never vanish simultaneously and are not all constant, and such that

$$\sum_{j=1}^{n} f_j^d=0.$$

Since $d\geq 25\geq (n-1)^2$, Lemma 7 shows that two of the terms must be proportional.

We now define an equivalence relation by taking $i\sim j$ if f_i and f_j are proportional. Repeated application of Lemma 7 then shows that the sum of the terms from each equivalence class must vanish. If any equivalence class has size 3 we deduce that, in terms of the original variables, all relevant points on C satisfy a pair of trinomial equations of the form

$$x_i^d+x_j^d=x_k^d.$$

Here we recall that only solutions of (1.1) in positive integers are to be considered. Wiles' proof [15] of Fermat's Theorem shows that one variable from each such equation must be zero, and we then see that only the trivial solutions of (1.1) can arise in this way.

If no equivalence class has size 3 we see that there are indices $i<j$ such that all relevant points on C satisfy either $x_i^d+x_j^d=0$ (if $i,j\leq 3$ or $i,j\geq 4$) or $x_i^d=x_j^d$ (if $i\leq 3<j$). Any point $\mathbf{x}$ with $x_i\in\mathbb{N}$, as required for our theorem, must therefore lie on one of the hyperplanes $x_i=x_j$ for $i\leq 3<j$. Taking $i=3, j=6$ for the sake of argument, we have therefore to consider solutions in positive integers of the equation

$$x_1^d+x_2^d=x_4^d+x_5^d.$$

This has been studied by a number of authors, but for our purposes an old result given by Greaves [4] suffices. This states that, for $d\geq 4$, there are $O_\varepsilon(B^{11/6+\varepsilon})$ non-trivial solutions of height at most B. Since there are $O(B)$ choices for the remaining variables $x_3=x_6$, we see that points lying on curves of degree at most 3 can contribute $O_\varepsilon(B^{17/6+\varepsilon})$ to $N_d^{(0)}(B)$. We state this formally as follows.

Lemma 8 *Let $d\geq 25$. Then for any $\varepsilon>0$, the contribution to $N_d^{(0)}(B)$ arising from points which lie on curves of degree at most 3, contained in the hypersurface (1.1), is $O_\varepsilon(B^{17/6+\varepsilon})$.*

5 Proof of the Theorem

We remark at the outset of the proof that it is sufficient to count primitive integer points. In general, if we know there are $O(B^{\phi})$ primitive points, for some constant $\phi > 1$, there will be $O(\sum_{n \ll B}(B/n)^{\phi}) = O(B^{\phi})$ points in total.

For any vector $\mathbf{x} = (x_1, \ldots, x_6) \in \mathbb{Z}^6$ we shall write $\hat{\mathbf{x}} = (x_1, \ldots, x_5)$. Thus if $\mathbf{x}$ is a primitive solution of $F(\mathbf{x}) = 0$ then $\hat{\mathbf{x}}$ must be non-zero, and will also be primitive. Then the lattice $\{\mathbf{y} \in \mathbb{Z}^5 : \hat{\mathbf{x}}.\mathbf{y} = 0\}$ has dimension 4. Moreover it is primitive and has determinant

$$|\hat{\mathbf{x}}| \ll |\mathbf{x}| \ll B,$$

by part (i) of Lemma 1. According to Lemma 1, part (ii), there is a basis $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(4)}$ of the lattice, for which

$$|\mathbf{y}^{(1)}| \leq \ldots \leq |\mathbf{y}^{(4)}|$$

and

$$\prod |\mathbf{y}^{(i)}| \ll |\mathbf{x}| \ll B.$$

We do not claim that the basis $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(4)}$ is unique, but we fix such a basis for each $\mathbf{x}$. We then define $S(\mathbf{y}^{(1)}, \mathbf{y}^{(2)})$ to be the set of vectors $\mathbf{x}$ for which the first two basis vectors are $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}$. We divide the available range for the $|\mathbf{y}^{(i)}|$ into dyadic intervals $Y_i < |\mathbf{y}^{(i)}| \leq 2Y_i$, so that

$$1 \ll Y_1 \ll \ldots \ll Y_4 \quad \text{and} \quad Y_1Y_2Y_3Y_4 \ll B. \tag{5.1}$$

It therefore follows that

$$Y_1^a Y_2^b \ll B^{(a+b)/4}, \tag{5.2}$$

whenever $b \leq 3a$. We will find it notationally convenient to write $\mathbf{y}^{(1)} = \mathbf{a}$ and $\mathbf{y}^{(2)} = \mathbf{b}$. Our overall strategy is to count points in $S(\mathbf{a}, \mathbf{b})$, all of which lie in the intersection

$$F(\mathbf{x}) = \mathbf{a}.\hat{\mathbf{x}} = \mathbf{b}.\hat{\mathbf{x}} = 0.$$

We henceforth let $F_{\mathbf{a}}$ denote the threefold $F(\mathbf{x}) = \mathbf{a}.\hat{\mathbf{x}} = 0$, and let $F_{\mathbf{a},\mathbf{b}}$ denote the corresponding surface given above. Our estimates for $\#S(\mathbf{a}, \mathbf{b})$ will depend on whether $F_{\mathbf{a},\mathbf{b}}$ is non-singular, or is irreducible but singular, or is reducible.

We begin by considering those vectors $\mathbf{x}$ for which $F_{\mathbf{a},\mathbf{b}}$ is irreducible, and shall use the fact that for each point on the surface the vector $\hat{\mathbf{x}}$ lies in the 3-dimensional lattice

$$\mathsf{M} = \{\mathbf{w} \in \mathbb{Z}^5 : \mathbf{a}.\mathbf{w} = \mathbf{b}.\mathbf{w} = 0\}.$$

According to part (iv) of Lemma 1, the determinant of M satisfies

$$Y_1Y_2 \ll \det(\mathsf{M}) \ll Y_1Y_2.$$

Moreover there will be a basis $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{u}^{(3)}$ of M satisfying part (ii) of Lemma 1 . In particular (2.3) takes the shape

$$Y_1Y_2 \ll |\mathbf{u}^{(1)}||\mathbf{u}^{(2)}||\mathbf{u}^{(3)}| \ll Y_1Y_2,$$

and if $\hat{\mathbf{x}} \in \mathsf{M}$ has $|x_i| \le B$, then (2.2) implies that

$$\hat{\mathbf{x}} = \lambda_1 \mathbf{u}^{(1)} + \lambda_2 \mathbf{u}^{(2)} + \lambda_3 \mathbf{u}^{(3)} \tag{5.3}$$

with

$$\lambda_i \ll B/|\mathbf{u}^{(i)}| = B_i \quad (i = 1, 2, 3). \tag{5.4}$$

Here, $B_1 \gg B_2 \gg B_3 \gg B^{1/2}$, since $|\mathbf{u}^{(3)}| \ll Y_1 Y_2 \ll B^{1/2}$ by (5.2). As before, we do not claim that $\mathbf{a}$ and $\mathbf{b}$ determine $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{u}^{(3)}$ uniquely, but we associate a particular choice of these vectors to every pair $\mathbf{a}, \mathbf{b}$.

We observe that under the substitution (5.3), $F_{\mathbf{a},\mathbf{b}}$ is given by an equation of the shape

$$f(\lambda_1, \lambda_2, \lambda_3) + x_6^d = 0, \tag{5.5}$$

where f is an appropriate ternary form of degree d defined over $\mathbb{Q}$. In order to count points on $F_{\mathbf{a},\mathbf{b}}$, we shall use Lemma 3 to reduce the analysis to that of a small number of curves within the surface. First we note via (5.4) and (5.5) that we can take $T = B^d$ in the statement of Lemma 3. Moreover we have

$$B_1 B_2 B_3 \ll \frac{B^3}{|\mathbf{u}^{(1)}|.|\mathbf{u}^{(2)}|.|\mathbf{u}^{(3)}|} \ll \frac{B^3}{Y_1 Y_2}$$

and $\log \|F_{\mathbf{a},\mathbf{b}}\| \ll \log B$. Thus Lemma 3 shows that the number of curves needed is

$$\ll_\varepsilon \frac{B^{3/\sqrt{d}+\varepsilon}}{(Y_1 Y_2)^{1/\sqrt{d}}}. \tag{5.6}$$

Moreover Lemma 5, in conjunction with (5.4), shows that each irreducible curve of degree $\delta \ll 1$ contributes

$$\ll_\varepsilon B^{1/\delta+\varepsilon} B_2^{1/\delta} = B^{2/\delta+\varepsilon} |\mathbf{u}^{(2)}|^{-1/\delta}. \tag{5.7}$$

It will be necessary to gain a saving from the factor $|\mathbf{u}^{(2)}|^{-1/\delta}$ in (5.7), by showing that $\mathbf{u}^{(2)}$ cannot be too small on average. Suppose that

$$C_j < |\mathbf{u}^{(j)}| \le 2C_j,$$

so that in particular $1 \ll C_1 \ll C_2 \ll C_3$ and $Y_1 Y_2 \ll C_1 C_2 C_3 \ll Y_1 Y_2$. Let

$$N(Y_1, Y_2, C_2) = \#\{\mathbf{a}, \mathbf{b} : \ Y_1 < |\mathbf{a}| \le 2Y_1,\ Y_2 < |\mathbf{b}| \le 2Y_2,\ C_2 < |\mathbf{u}^{(2)}| \le 2C_2\}.$$

In order to estimate $N(Y_1, Y_2, C_2)$ we use the fact that both $\mathbf{a}$ and $\mathbf{b}$ belong to the 4-dimensional lattice

$$\mathsf{L} = \{\mathbf{w} \in \mathbb{Z}^5 : \mathbf{w}.\mathbf{u}^{(2)} = 0\}.$$

By part (i) of Lemma 1, the determinant of L has order of magnitude C_2. Moreover part (iii) of Lemma 1 then shows that there can be at most $O(Y_2^4/C_2)$ values of $\mathbf{b}$ lying in L, since $C_2 \ll (Y_1 Y_2)^{1/2} \ll Y_2$. Similarly we conclude that there are $O(Y_1^4/C_2 + Y_1^3)$ values of $\mathbf{a}$ lying in L. Summing up over the $O(C_2^5)$ values of $\mathbf{u}^{(2)}$, we therefore deduce that

$$N(Y_1, Y_2, C_2) \ll C_2^4(Y_1^4/C_2 + Y_1^3)Y_2^4. \tag{5.8}$$

Alternatively we may of course use the trivial bound

$$N(Y_1, Y_2, C_2) \ll (Y_1Y_2)^5. \tag{5.9}$$

We use the first of these estimates when $C_2 \le B^{1/6}$ and the second when $C_2 > B^{1/6}$. Then (5.1) implies that $Y_2 \ll B^{1/3}$, whence (5.2) leads to

$$\begin{aligned} &B^{2/\delta+\varepsilon}C_2^{-1/\delta}N(Y_1, Y_2, C_2) \\ &\quad \ll B^{11/6\delta+\varepsilon}\{(Y_1Y_2)^4B^{1/2} + Y_1^3Y_2^4B^{2/3} + (Y_1Y_2)^5\} \\ &\quad \ll B^{11/6\delta+\varepsilon}(Y_1Y_2)^3B. \end{aligned} \tag{5.10}$$

We now split our consideration of the case in which $F_{\mathbf{a},\mathbf{b}}$ is irreducible into two further sub-cases, starting with the sub-case in which $F_{\mathbf{a},\mathbf{b}}$ is non-singular. Here we begin by examining curves of degree $\le d-2$ contained in the surface. For each relevant pair $\mathbf{a}, \mathbf{b}$ the number of these is bounded in terms of d alone by a result of Colliot-Thélène [8; Appendix]. Moreover points that lie on curves of degree at most 3 have already been dealt with in Lemma 8. Thus the total contribution to $N_d^{(0)}(B)$ arising from curves with degree $\delta \le d-2$, and

$$Y_1 < |\mathbf{a}| \le 2Y_1, \quad Y_2 < |\mathbf{b}| \le 2Y_2$$

is given by (5.10) with $\delta \ge 4$. This produces a satisfactory overall estimate

$$\ll_\varepsilon B^{3-1/24+\varepsilon}$$

by (5.1) On the other hand, for curves of degree $\delta \ge d-1$ we have the bound (5.6) for the number of curves corresponding to a particular pair $\mathbf{a}, \mathbf{b}$. Thus (5.10) easily leads to the overall contribution

$$\ll_\varepsilon B^{5/2+5/2\sqrt{d}+11/6(d-1)+\varepsilon}$$

from this case. Summing the C_j over dyadic ranges, we therefore complete the estimation of the number of points lying on non-singular surfaces $F_{\mathbf{a},\mathbf{b}}$.

The next case we shall consider is that in which $F_{\mathbf{a}}$ is non-singular, but $F_{\mathbf{a},\mathbf{b}}$ is singular and irreducible. To produce a convenient model for the threefold $F_{\mathbf{a}}$ we shall suppose that $a_1 \ne 0$, say. We may then eliminate x_1 via the equation $\mathbf{a}.\hat{\mathbf{x}} = 0$, to produce a model

$$f_{\mathbf{a}}(x_2, x_3, x_4, x_5) + a_1^d x_6^d = 0$$

for $F_{\mathbf{a}}$. Clearly the non-singularity of $F_{\mathbf{a}}$ entails the non-singularity of the surface $f_{\mathbf{a}} = 0$. In particular there exists a non-zero dual form

$$\hat{f}_{\mathbf{a}}(\mathbf{t}) \in \mathbb{Q}[t_1, t_2, t_3, t_4],$$

which vanishes whenever the plane section $f_{\mathbf{a}}(\mathbf{X}) = \mathbf{t}.\mathbf{X} = 0$ produces a singular curve. Furthermore, $\hat{f}_{\mathbf{a}}$ is irreducible and has degree $2 \le \delta \ll 1$.

We can also use the relation $\mathbf{a}.\hat{\mathbf{x}} = 0$ to eliminate x_1 from the equation $\mathbf{b}.\hat{\mathbf{x}} = 0$. In this way we find that $F_{\mathbf{a},\mathbf{b}}$ will have a model

$$f_{\mathbf{a}}(x_2, x_3, x_4, x_5) + a_1^d x_6^d = c_2 x_2 + c x_3 x_3 + c_4 x_4 + c_5 x_5 = 0,$$

where

$$c_j = a_1 b_j - a_j b_1 \quad \text{for} \quad j = 2, 3, 4, 5.$$

It is now clear that $F_{\mathbf{a},\mathbf{b}}$ is singular if and only if the intersection

$$f_{\mathbf{a}}(x_2, x_3, x_4, x_5) = c_2 x_2 + c x_3 x_3 + c_4 x_4 + c_5 x_5 = 0$$

is singular, and this holds if and only if $\hat{f}_{\mathbf{a}}(\mathbf{c}) = 0$. We now set

$$g(\mathbf{b}) = \hat{f}_{\mathbf{a}}(a_1 b_2 - a_2 b_1, \ldots, a_1 b_5 - a_5 b_1),$$

and consider points $\mathbf{b}$ lying on the threefold $g(\mathbf{b}) = 0$. Since $g(0, b_2, b_3, b_4, b_5)$ is irreducible of degree δ we conclude that $g(\mathbf{b})$ is also irreducible with degree δ. Thus we may apply Lemma 4 to deduce that for each $\mathbf{a}$ there are at most $O_\varepsilon(Y_2^{10/3+\varepsilon})$ vectors $\mathbf{b}$ which lead to singular $F_{\mathbf{a},\mathbf{b}}$. This leads to the overall bound $O_\varepsilon(Y_1^5 Y_2^{10/3+\varepsilon})$ for the number of pairs $\mathbf{a}, \mathbf{b}$ for which $F_{\mathbf{a}}$ is non-singular and $F_{\mathbf{a},\mathbf{b}}$ is singular.

We are now ready to return to the bound (5.6) for the number of curves that we must examine for a given pair $\mathbf{a}, \mathbf{b}$. In view of Lemma 8 we need only consider curves of degree $\delta \ge 4$, and according to (5.7) these contribute $O(B^{1/2+\varepsilon}|\mathbf{u}^{(2)}|^{-1/4})$ each. As before we shall need to estimate the analogue of $N(Y_1, Y_2, C_2)$ for the current situation. We therefore write $N'(Y_1, Y_2, C_2)$ for the number of pairs of vectors $\mathbf{a}, \mathbf{b}$ counted by the original function $N(Y_1, Y_2, C_2)$, with the property that $F_{\mathbf{a}}$ is non-singular and $F_{\mathbf{a},\mathbf{b}}$ is singular. Such pairs then contribute

$$\ll_\varepsilon \frac{B^{3/\sqrt{d}}}{(Y_1 Y_2)^{1/\sqrt{d}}} B^{1/2+\varepsilon} C_2^{-1/4} N'(Y_1, Y_2, C_2) \tag{5.11}$$

to $N_d^{(0)}(B)$. We have shown above that (5.9) may be replaced by

$$N'(Y_1, Y_2, C_2) \ll_\varepsilon Y_1^5 Y_2^{10/3+\varepsilon}.$$

On the other hand, if we specify $\mathbf{u}^{(2)}$ then there are $O(Y_1^4/C_2 + Y_1^3)$ choices for $\mathbf{a}$. To count the number of available vectors $\mathbf{b}$ we use the equations $g(\mathbf{b}) = \mathbf{b}.\mathbf{u}^{(2)} = 0$ to eliminate one of the variables b_5, say, producing a condition $G(b_1, b_2, b_3, b_4) = 0$. The form G cannot vanish identically since $g(\mathbf{t})$ is not divisible by the linear form $\mathbf{t}.\mathbf{u}^{(2)}$. We are therefore in a position to apply Lemma 2, which shows that we have $O(Y_2^3)$ vectors $\mathbf{b}$, once $\mathbf{u}^{(2)}$ and $\mathbf{a}$ have been chosen. This leads to the alternative bound

$$N'(Y_1, Y_2, C_2) \ll C_2^5 (Y_1^4/C_2 + Y_1^3) Y_2^3,$$

which we use as a substitute for (5.8). We now proceed as before, using the first estimate for $N'(Y_1,Y_2,C_2)$ when $C_2\geq B^{1/12}$ and the second otherwise. These bounds produce

$$C_2^{-1/4}N'(Y_1,Y_2,C_2)\ll_\varepsilon (Y_1^2Y_2^{1/3+\varepsilon}B^{-1/48}+Y_1B^{5/16}+B^{19/48})(Y_1Y_2)^3,$$

so that the estimate (5.11) becomes

$$\begin{aligned}
&\ll_\varepsilon \quad \frac{B^{3/\sqrt{d}}}{(Y_1Y_2)^{1/\sqrt{d}}}B^{1/2+\varepsilon}(Y_1^2Y_2^{1/3+\varepsilon}B^{-1/48}+Y_1B^{5/16}+B^{19/48})(Y_1Y_2)^3\\
&\ll_\varepsilon \quad B^{41/16+5/2\sqrt{d}}
\end{aligned}$$

in view of (5.2). This is satisfactory for our theorem.

We have now completely disposed of the cases in which either $F_{\mathbf{a},\mathbf{b}}$ is non-singular, or $F_{\mathbf{a}}$ is non-singular and $F_{\mathbf{a},\mathbf{b}}$ is singular but irreducible. We turn now to the case in which $F_{\mathbf{a}}$ is singular, and $F_{\mathbf{a},\mathbf{b}}$ is singular but irreducible. According to Lemma 6 there are at most $O(Y_1^2)$ possible vectors $\mathbf{a}$, for each of the $O(Y_2^5)$ choices for $\mathbf{b}$. Taking the lower bound $Y_1Y_2\gg 1$ in (5.6), together with the lower bound $D\geq 4$ in Lemma 5, we deduce that each surface $F_{\mathbf{a},\mathbf{b}}$ contributes $O_\varepsilon(B^{1/2+3/\sqrt{d}+\varepsilon})$ points, apart from any that lie on curves of degree at most 3. Thus we obtain the overall contribution

$$\ll_\varepsilon B^{9/4+3/\sqrt{d}+\varepsilon},$$

since $Y_1^2Y_2^5\ll B^{7/4}$ by (5.2). This is also satisfactory, since $d\geq 25$ and the points on curves of degree at most 3 are catered for by Lemma 8.

Finally we must handle the situation in which $F_{\mathbf{a},\mathbf{b}}$ is reducible. In this case it follows that we can write

$$F(\mathbf{x})=(\mathbf{a}.\hat{\mathbf{x}})R(\mathbf{x})+(\mathbf{b}.\hat{\mathbf{x}})S(\mathbf{x})+U(\mathbf{x})V(\mathbf{x})$$

for appropriate forms R,S,U,V. On eliminating one variable via the equation $\mathbf{a}.\hat{\mathbf{x}}=0$ we can think of $F_{\mathbf{a}}$ as being given by $(\mathbf{b}.\hat{\mathbf{x}})S(\mathbf{x})+U(\mathbf{x})V(\mathbf{x})$. However the 5 simultaneous conditions

$$\mathbf{a}.\hat{\mathbf{x}}=\mathbf{b}.\hat{\mathbf{x}}=S(\mathbf{x})=U(\mathbf{x})=V(\mathbf{x})=0$$

in the 6 variables x_i necessarily have a non-zero solution, which will be a singular point of $F_{\mathbf{a}}$. We conclude that if $F_{\mathbf{a},\mathbf{b}}$ is reducible, then $F_{\mathbf{a}}$ is singular, and similarly for $F_{\mathbf{b}}$. Thus the only case left to dispose of is that in which both $F_{\mathbf{a}}$ and $F_{\mathbf{b}}$ are singular, and Lemma 6 shows that this situation can arise for at most $O(Y_1^2Y_2^2)$ values of $\mathbf{a},\mathbf{b}$. Let G be an irreducible component of $F_{\mathbf{a},\mathbf{b}}$ of degree $\delta\leq d$, and suppose firstly that $\delta\leq 3$. For any rational point $\mathbf{x}$ on the surface G take a rational plane H containing $\mathbf{x}$ but not containing the entire surface G. Then $\mathbf{x}$ lies on some irreducible component C of $S\cap H$. Moreover C is a curve of degree at most 3 and is contained in the original hypersurface $F=0$. Thus $\mathbf{x}$ has already been allowed for under Lemma 8. It therefore suffices to sum the contribution from the points lying on those components G for which $\delta \geq 4$. Points which lie on lines contained in the surface G can be neglected, by a further application on Lemma 8. We may therefore apply the second author's bound [8; Theorem 6] for the number of non-trivial points on an irreducible surface. This shows that G contains $O_\varepsilon(B^{52/27+\varepsilon})$ points not lying on rational lines in the surface, which therefore leads to an overall contribution

$$\ll_\varepsilon (Y_1Y_2)^2 B^{52/27+\varepsilon} \ll_\varepsilon B^{79/27+\varepsilon}$$

to $N_d^{(0)}(B)$. This is clearly satisfactory for our theorem, and completes the proof.

6 Proof of the Corollary

It remains to establish the corollary. The formula (1.2) is trivial, since the sum is merely

$$\sum_{m^d+n^d\leq x} [(x-m^d-n^d)^{1/d}],$$

which one may approximate by the corresponding integral. The sum in the corollary counts solutions of

$$n_1^d + n_2^d + n_3^d = n_4^d + n_5^d + n_6^d \leq x$$

in positive integers n_i. Any solution in which n_4, n_5, n_6 are not a permutation of n_1, n_2, n_3 will be counted by $N_d^{(0)}(x^{1/d})$, so there will be $o(x^{3/d})$ such solutions, for $d \geq 33$. Amongst the trivial solutions, there will be $O(x^{2/d})$ in which n_1, n_2, n_3 are not distinct, whence

$$\sum_{n\leq x} r(n)^2 = 6\sum_{n\leq x} r(n) + o(x^{3/d}) = 6cx^{3/d} + o(x^{3/d}),$$

as required. For the second assertion in the corollary, we take $N(x;d)$ to be the number of positive integers $n \leq x$ which are sums of 3 positive d-th powers, and $N^{(*)}(x;d)$ to be the number of such integers with two or more essentially distinct representations. Write $r_0(n) = 1$ if n is a sum of 3 positive d-th powers, and $r_0(n) = 0$ otherwise, so that

$$N(x;d) = \sum_{n\leq x} r_0(n).$$

With the exception of $O(x^{2/d})$ integers $n \leq x$, we have $r(n) \geq 6$ whenever $r_0(n) = 1$. Hence

$$6N(x;d) = 6\sum_{n\leq x} r_0(n) \leq \sum_{n\leq x} r(n) + O(x^{2/d}) = cx^{3/d} + o(x^{3/d}). \tag{6.1}$$

On the other hand Cauchy's inequality yields

$$\{\sum_{n\leq x} r(n)\}^2 \leq N(x;d)\sum_{n\leq x} r(n)^2,$$

whence

$$\{cx^{3/d}+o(x^{3/d})\}^2 \leq N(x;d)\{6cx^{3/d}+o(x^{3/d})\}.$$

It follows that $N(x;d) \geq \frac{1}{6}cx^{3/d}+o(x^{3/d})$. In view of (6.1) we therefore have $N(x;d) \sim \frac{1}{6}cx^{3/d}$. Finally we note that, if $n \leq x$ has two or more essentially distinct representations as sums of 3 positive d-th powers, then, with the exception of $O(x^{2/d})$ possible integers n, we will have $r(n) > 6r_0(n)$. Thus

$$\begin{aligned} N^{(*)}(x;d) &\leq \sum_{n\leq x}(r(n)-6r_0(n))+O(x^{2/d}) \\ &= \{cx^{3/d}+o(x^{3/d})\}-6\{\frac{1}{6}cx^{3/d}+o(x^{3/d})\}+O(x^{2/d}) \\ &= o(x^{3/d}). \end{aligned}$$

It follows that almost all integers which are sums of 3 positive d-th powers have essentially one such representation.

References

[1] V. Batyrev and Yu. I. Manin, Sur le nombre des points rationnels de hauteur borné des variétés algébriques, *Math. Ann.,* 286 (1990), 27–43.

[2] H. Davenport, Cubic forms in 16 variables, *Proc. Roy. Soc. A,* 272 (1963), 285–303.

[3] L.E. Dickson, *History of the theory of numbers,* (Chelsea, New York, 1966).

[4] G.R.H. Greaves, Representation of a number by the sum of two fourth powers, (Russian) *Mat. Zametki,* 55 (1994), 47–58. Translation in *Math. Notes.,* 55 (1994), 134–141.

[5] M.L. Green, Some Picard theorems for holomorphic maps to algebraic varieties, *Amer. J. Math.,* 97 (1975), 43–75.

[6] D.R. Heath-Brown, Diophantine approximation with square-free numbers, *Math. Zeit.,* 187 (1984), 335–344.

[7] D.R. Heath-Brown, The circle method and diagonal cubic forms, *Phil. Trans. Roy. Soc. London,* A 356 (1998), 673–699.

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

[9] C. Hooley, On some topics connected with Waring's problem, *J. Reine Angew. Math.,* 369 (1986), 110–153.

[10] C. Hooley, On Hypothesis K^* in Waring's problem, *Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995),* 175–185, London Math. Soc. Lecture Note Ser., 237, (Cambridge Univ. Press, Cambridge, 1997).

[11] L.K. Hua, On Waring's problem, *Quart. J. Math. Oxford Ser.,* 9 (1938), 199–202.

[12] D.J. Newman and M. Slater, Waring's problem for the ring of polynomials, *J. Number Theory,* 11 (1979), 477–487.

[13] J. Pila, Density of integral and rational points on varieties, *Astérisque,* 228 (1995), 183–187.

[14] R.C. Vaughan, and T.D. Wooley, Further improvements in Waring's problem, *Acta Math.,* 174 (1995), 147–240.

[15] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2),* 141 (1995), 443–551.

[16] T.D. Wooley, Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour, *Invent. Math.,* 122 (1995), 421–51.