

# Counting Rational Points on Hypersurfaces

T.D. Browning<sup>1</sup>

D.R. Heath-Brown<sup>2</sup>

*Mathematical Institute, 24–29 St. Giles', Oxford OX1 3LB*

<sup>1</sup>browning@maths.ox.ac.uk, <sup>2</sup>rhb@maths.ox.ac.uk

## Abstract

For any  $n \geq 2$ , let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a form of degree  $d \geq 2$ , which produces a geometrically irreducible hypersurface in  $\mathbb{P}^{n-1}$ . This paper is concerned with the number  $N(F; B)$  of rational points on  $F = 0$  which have height at most  $B$ . For any  $\varepsilon > 0$  we establish the estimate

$$N(F; B) = O(B^{n-2+\varepsilon}),$$

whenever either  $n \leq 5$  or the hypersurface is not a union of lines. Here the implied constant depends at most upon  $d, n$  and  $\varepsilon$ .

## 1 Introduction

For any  $n \geq 2$ , let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a non-zero form of degree  $d$ , which produces an algebraic hypersurface  $X$  in  $\mathbb{P}^{n-1}$ . Our basic interest is with the distribution of rational points on such hypersurfaces. It will be convenient to write  $Z^n$  for the set of primitive vectors  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ , where  $\mathbf{x}$  is said to be primitive if  $\text{h.c.f.}(x_1, \dots, x_n) = 1$ . With this notation, we seek to understand the asymptotic behaviour of the quantity

$$N(F; B) = \#\{\mathbf{x} \in Z^n : F(\mathbf{x}) = 0, \max_i |x_i| \leq B\},$$

as  $B \rightarrow \infty$ . More precisely, this paper is motivated by the following conjecture of the second author [9, Conjecture 2].

**Conjecture.** *Let  $\varepsilon > 0$ . Suppose that  $F$  is irreducible and that  $d \geq 2$ . Then we have*

$$N(F; B) = O(B^{n-2+\varepsilon}).$$

Here, and throughout this paper, the implied constant may depend at most upon  $d, n$  and the choice of  $\varepsilon$ . Any further dependences will be explicitly indicated. In the statement of the above conjecture, and elsewhere, we shall always take irreducibility of a form to mean absolute irreducibility. One might also ask about bounds of the shape  $N(F; B) = O_F(B^{n-2+\varepsilon})$ , as in [9, Conjecture 1], where the implied constant is allowed to depend on the coefficients of  $F$ . However it transpires that estimates which are uniform in forms of fixed degree, and with a fixed number of variables, are much more useful in applications, and that in most cases results with weaker uniformity appear to be no easier to prove.

We take a moment to discuss the available evidence for the conjecture. Firstly it should be clear that the bound's exponent is in general essentially as sharp as can be hoped for. Indeed, whenever the hypersurface  $X$  contains a linear space defined over  $\mathbb{Q}$  and having dimension  $n - 3$ , then we automatically have  $N(F; B) \gg_F B^{n-2}$ . The second author has already established the conjecture in the case of quadrics [9, Theorem 2], in addition to the cases  $n = 3$  and  $n = 4$  for any degree [9, Theorems 3 and 9]. More recently, Broberg and Salberger [2] have examined the case  $n = 5$  of threefolds. They succeed in establishing the conjecture as soon as  $d \geq 4$ . The best result available in most other cases is the bound

$$N(F; B) = O(B^{n-2+1/d+\varepsilon}), \quad (1.1)$$

due to Pila [10]. In the case of cubic threefolds, Broberg and Salberger [2] have improved upon the exponent  $3 + 1/3$  appearing in (1.1), allowing one to replace it by  $3 + 1/18$ .

Before proceeding further we record the following rather general “trivial” bound, in which  $[\mathbf{x}] \in \mathbb{P}^{N-1}$  denotes the projective point corresponding to a vector  $\mathbf{x} \in \mathbb{C}^N$  for any  $N \geq 2$ . This will be established in §2.

**Theorem 1.** *Let  $Y \subseteq \mathbb{P}^{N-1}$  be an irreducible variety defined over  $\overline{\mathbb{Q}}$ , of dimension  $m$  and degree  $D$ . Then*

$$\#\{\mathbf{x} \in \mathbb{Z}^N : [\mathbf{x}] \in Y, \max_i |x_i| \leq B\} \ll_{D,N} B^{m+1}.$$

Here, and elsewhere, when we say that a variety is irreducible, we shall mean that it is geometrically irreducible. We may conclude from Theorem 1 that points which lie on a subvariety of  $X$  of degree  $O(1)$ , and codimension 1 or more, make an acceptable contribution for our conjecture. The estimate is clearly best possible in the case of linear varieties.

Our first substantial result shows that it is those points that lie on lines in the hypersurface  $X$  which are the only real difficulty. In fact we shall be able to tackle a rather more general counting problem in which we consider points in a box, rather than a cube. To describe this situation we let  $B_i \geq 1$  for  $1 \leq i \leq n$ , and write  $\mathbf{B} = (B_1, \dots, B_n)$  and

$$V = \prod_{i=1}^n B_i. \quad (1.2)$$

We then define

$$N(F; \mathbf{B}) = \#\{\mathbf{x} \in \mathbb{Z}^n : F(\mathbf{x}) = 0, |x_i| \leq B_i, (1 \leq i \leq n)\}.$$

In particular we clearly have  $N(F; \mathbf{B}) = N(F; B)$  in the notation above, whenever  $B_i = B$  for  $1 \leq i \leq n$ . Moreover we shall define  $N_1(F; \mathbf{B})$  to be the number of points counted by  $N(F; \mathbf{B})$ , but which do not lie on any line contained in  $X$ . In the special case  $B_i = B$  for  $1 \leq i \leq n$ , we merely write  $N_1(F; B)$  for  $N_1(F; \mathbf{B})$ . Finally we write  $\|F\|$  for the maximum modulus of the coefficients of  $F$ . We then have the following two results, which will be established in §3.

**Theorem 2.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be an irreducible form of degree  $d$ . Then we have*

$$N_1(F; \mathbf{B}) \ll V^{(n-2)/n} (\log \|F\| V)^{(n-2)(2n^2+n+3)/3},$$

and

$$N_1(F; B) = O(B^{n-2+\varepsilon}).$$

**Corollary 1.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be an irreducible form of degree  $d$ , and suppose that the variety  $X$  is not a union of lines. Then*

$$N(F; B) = O(B^{n-2+\varepsilon}).$$

Handling the lines on the variety  $X$  requires considerable further work, and we have only been successful in extending the results proved in [9] for  $n = 3$  and  $n = 4$ , to the case  $n = 5$  corresponding to threefolds. We shall establish the following result, which provides further evidence for the main conjecture.

**Theorem 3.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_5]$  be an irreducible form of degree  $d \geq 2$ . Then we have*

$$N(F; B) = O(B^{3+\varepsilon}).$$

In fact our attack on the conjecture shows that every point counted by  $N(F; \mathbf{B})$  must lie on one of a small number of linear subspaces of the hypersurface  $X$ . The contribution from the points lying on linear spaces of dimension 0 (that is to say, from individual points) will be satisfactory from the point of view of the conjecture, while those lying on linear spaces of higher dimension will be problematic. In order to state the result we must first introduce some more notation. Let  $\Lambda \subseteq \mathbb{Z}^n$  be any integer lattice of dimension  $m \geq 2$  and determinant  $\det \Lambda$ , and define the set

$$S(F; B, \Lambda) = \{\mathbf{x} \in \Lambda \cap \mathbb{Z}^n : F(\mathbf{x}) = 0, \max_i |x_i| \leq B\}.$$

Let  $s_1 \leq \dots \leq s_m$  be the successive minima of  $\Lambda$  with respect to the Euclidean length. A brief discussion of successive minima will be included in §2. We then have the following result.

**Theorem 4.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a form of degree  $d$ , not necessarily irreducible, but not vanishing identically on  $\Lambda$ . Let  $p_0$  be the smallest prime  $p_0 > n$ , and let  $B \geq s_m$ . Then there exist lattices  $\mathbf{M}_1, \dots, \mathbf{M}_J \subseteq \Lambda$ , such that the linear space  $M_j$  spanned by each  $\mathbf{M}_j$  lies in the variety  $X$ , and such that*

$$J \ll \left( \frac{B^m}{\det \Lambda} \right)^{(m-2)/m} (\log \|F\| B)^{(m-2)(2m^2+m+3)/3}. \quad (1.3)$$

Moreover the successive minima of  $\mathbf{M}_j$  are all  $O(B(\log \|F\| B)^{m(m+1)})$  and

$$S(F; B, \Lambda) \subseteq \bigcup_{j=1}^J \mathbf{M}_j.$$

Associated to each proper sublattice  $\mathbf{M}_j \subset \Lambda$  of dimension  $m - h$ , are positive integers  $q_1, \dots, q_h$  and a vector  $\mathbf{w} \in \mathbb{Z}^n$ , all depending on  $\mathbf{M}_j$ , and such that  $q_i$  is a power of a prime  $p_i \equiv m + 1 - i \pmod{p_0}$ . These have the property that

$$\mathbf{M}_j \subseteq \{\mathbf{x} \in \Lambda : \mathbf{x} \equiv \rho \mathbf{w} \pmod{q_1 \cdots q_h} \text{ some } \rho \in \mathbb{Z}\}, \quad (1.4)$$

and

$$\det \mathbf{M}_j \ll (\det \Lambda) \prod_{i=1}^h \frac{q_i^{m-(i-1)}}{B}. \quad (1.5)$$

Evidently we will have  $\dim M_j = 1 + \dim M_j$ , since we are using the dimension of  $M_j$  as a projective space. The above result allows us to provide information about the heights of the linear spaces  $M_j$  as follows. This is crucial for the proof of Theorem 3. The height of a linear space will be defined in §7.

**Theorem 5.** *For any  $B \geq 1$  and  $n \geq 4$ , let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be an irreducible non-zero form of degree  $d$ . Let  $\Lambda = \mathbb{Z}^n$  in Theorem 4 and suppose that*

$$M_1, \dots, M_J \subseteq X$$

*are the resulting linear spaces, which cover the points of  $S(F; B, \mathbb{Z}^n)$ . Let  $1 \leq j \leq J$ . Then  $M_j$  is defined over  $\mathbb{Q}$ , and whenever the dimension of  $M_j$  is non-zero we have*

$$H(M_j) \ll B(\log \|F\|B)^{n^3}.$$

We stress that Theorem 4 refers to an arbitrary integer lattice  $\Lambda$ , while Theorem 5 considers only  $\mathbb{Z}^n$ . By change of basis, we shall see in §3 that there is an approximate equivalence between points of a general lattice in a cube, and points of  $\mathbb{Z}^n$  in an arbitrary box. However in order to estimate  $H(M_j)$  in Theorem 5 it is necessary to restrict attention to points of  $\mathbb{Z}^n$  in a cube.

It is interesting to see what Theorem 2 has to say about the quantity  $N(F; \mathbf{B})$  in the case of irreducible curves of degree  $d \geq 2$ . Taking  $n = 3$  we have

$$N(F; \mathbf{B}) \ll V^{1/3}(\log \|F\|V)^8.$$

This may be compared with work of the second author [9, Theorem 3], which implies that

$$N(F; \mathbf{B}) \ll V^{1/(d+1)+\varepsilon}. \quad (1.6)$$

To see this we suppose for the moment that  $B_1 \geq B_2 \geq B_3$ , and note that an irreducible ternary form must contain a monomial in which  $x_1$  appears explicitly and also a monomial in which  $x_3$  does not appear. Hence we may take

$$T \geq \max\{B_1 B_3^{d-1}, B_2^d\} \geq V^{d/(d+1)}$$

in the notation of [9, Equation (1.8) of Theorem 3], which suffices to establish (1.6). Thus it follows that a direct application of Theorem 2 is weaker than this earlier result, except possibly when  $d = 2$ . However, in this latter case we are able to provide the following cleaner result.

**Theorem 6.** *Let  $Q \in \mathbb{Z}[x_1, x_2, x_3]$  be a non-singular quadratic form. Then we have*

$$N(Q; \mathbf{B}) \ll V^{1/3}.$$

In fact it is easy to see that this is best possible whenever  $B_i = B$  for  $1 \leq i \leq 3$ . As a direct consequence of Theorem 6 we are able to deduce a further corollary, which sharpens [8, Theorem 2] and [9, Corollary 2].

**Corollary 2.** *Let  $Q \in \mathbb{Z}[x_1, x_2, x_3]$  be a non-singular quadratic form with matrix  $\mathbf{M}$ . Let  $\Delta = |\det \mathbf{M}|$  and write  $\Delta_0$  for the highest common factor of the  $2 \times 2$  minors of  $\mathbf{M}$ . Then*

$$N(Q; \mathbf{B}) \ll \left\{ 1 + \left( \frac{V \Delta_0^{3/2}}{\Delta} \right)^{1/3} \right\} d(\Delta).$$

This is superior to [8, Theorem 2] in three respects. Firstly we have replaced the exponent  $1/3 + \varepsilon$  in [9, Corollary 2] by  $1/3$ . This is a direct result of our Theorem 6. Secondly we have replaced  $\Delta_0^2$  by  $\Delta_0^{3/2}$ . However it is already explicit in the proof of [8, Theorem 2] that this is permissible. Finally we have replaced the function  $d_3(\Delta)$  by  $d(\Delta)$ . It is apparent from a closer inspection of the proof of [8, Theorem 2] that this too is allowable.

Theorem 2 can also be used to good effect in the case  $n = 4$  of surfaces. We assume without loss of generality that  $B_1 \geq B_2 \geq B_3 \geq B_4$ , and suppose that  $F \in \mathbb{Z}[x_1, \dots, x_4]$  is any form of degree  $d$ . Then Theorem 2 implies that

$$N_1(F; \mathbf{B}) \ll V^{1/2}(\log \|F\|V)^{26}.$$

It is natural to ask, as in the case of curves above, whether or not we might expect a similar upper bound to hold for the quantity  $N(F; \mathbf{B})$  if  $F$  is irreducible and has degree  $d \geq 2$ . However simple examples of the type

$$F = x_3F_1 + x_4F_2,$$

for suitable forms  $F_1, F_2 \in \mathbb{Z}[x_1, \dots, x_4]$  of degree  $d - 1$ , demonstrate that we may have  $N(F; \mathbf{B}) \gg B_1B_2$  whenever the surface  $F = 0$  contains the line  $x_3 = x_4 = 0$ . It is then a trivial matter to deduce that we only have  $V^{1/2} \geq B_1B_2$  when  $B_1 = B_2 = B_3 = B_4$ . These remarks show that the following result, which will be established in §8, is essentially best possible.

**Theorem 7.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_4]$  be an irreducible form of degree  $d \geq 2$  and suppose that  $B_1 \geq B_2 \geq B_3 \geq B_4$ . Then we have*

$$N(F; \mathbf{B}) \ll (B_1B_2)^{1+\varepsilon}.$$

**Acknowledgement.** While working on this paper, the first author was supported by EPSRC grant number GR/R93155/01.

## 2 Preliminaries

Our arguments will require a number of facts from the geometry of numbers, which it will be convenient to collect together in this section. Most of these results can be found in the second author's work [7, §2], for example.

Let  $\Lambda \subseteq \mathbb{Z}^n \subset \mathbb{R}^n$  be a lattice of dimension (or rank)  $r$ , and let  $\mathbf{b}_1, \dots, \mathbf{b}_r$  be any basis for  $\Lambda$ . Then we define the determinant  $\det \Lambda$  of  $\Lambda$ , to be the  $r$ -dimensional volume of the parallelepiped generated by  $\mathbf{b}_1, \dots, \mathbf{b}_r$ . This is independent of the choice of basis. In fact if  $\mathbf{M}$  denotes the  $n \times r$  matrix formed from the vectors  $\mathbf{b}_1, \dots, \mathbf{b}_r$ , then  $\mathbf{M}^T \mathbf{M}$  is a real positive symmetric matrix and we have

$$(\det \Lambda)^2 = \det(\mathbf{M}^T \mathbf{M}). \quad (2.1)$$

We say that  $\Lambda$  is primitive if it is not properly contained in any other  $r$ -dimensional sublattice of  $\mathbb{Z}^n$ .

In addition to these facts, we shall also make repeated use of the successive minima of a lattice. Let  $\Lambda \subseteq \mathbb{Z}^n$  be a lattice of dimension  $r$ , and let  $\langle \mathbf{a}_1, \dots, \mathbf{a}_r \rangle$  denote the  $\mathbb{Z}$ -linear span of any set of vectors  $\mathbf{a}_1, \dots, \mathbf{a}_r \in \mathbb{R}^n$ . Then we construct a minimal basis of  $\Lambda$  in the following manner. First let  $\mathbf{m}_1 \in \Lambda$  be

any non-zero vector for which the Euclidean length  $|\mathbf{m}_1|$  is least. Next let  $\mathbf{m}_2 \in \Lambda \setminus \langle \mathbf{m}_1 \rangle$  be any vector for which  $|\mathbf{m}_2|$  is least. Continuing in this way we obtain a basis  $\mathbf{m}_1, \dots, \mathbf{m}_r$  for  $\Lambda$  in which  $|\mathbf{m}_1| \leq \dots \leq |\mathbf{m}_r|$ . The successive minima of  $\Lambda$  with respect to the Euclidean length are merely the numbers  $s_i = |\mathbf{m}_i|$ , for  $1 \leq i \leq r$ . Then we have

$$\prod_{i=1}^r s_i \ll \det \Lambda \leq \prod_{i=1}^r s_i. \quad (2.2)$$

In fact the upper bound in (2.2) is a special case of the more general inequality  $\det \Lambda \leq \prod_{i=1}^r |\mathbf{b}_i|$ , which holds for any basis  $\mathbf{b}_1, \dots, \mathbf{b}_r$  for  $\Lambda$ . A fundamental property of successive minima and minimal bases is recorded in the following result, for which see Davenport [5, Lemma 5].

**Lemma 1.** *Let  $\Lambda \subseteq \mathbb{Z}^n$  be a lattice of dimension  $r$ , with successive minima  $s_1 \leq \dots \leq s_r$ . Then  $\Lambda$  has a basis  $\mathbf{m}_1, \dots, \mathbf{m}_r$  such that  $|\mathbf{m}_i| = s_i$  for  $1 \leq i \leq r$ , and with the property that whenever one writes  $\mathbf{x} \in \Lambda$  as*

$$\mathbf{x} = \sum_{i=1}^r \lambda_i \mathbf{m}_i,$$

*then  $\lambda_i \ll |\mathbf{x}|/s_i$ , for  $1 \leq i \leq r$ .*

We shall also need a result which describes how the successive minima of a lattice relate to the successive minima of any sublattice.

**Lemma 2.** *Let  $\Lambda' \subseteq \Lambda \subseteq \mathbb{Z}^n$  be lattices of dimension  $r$ . Suppose that  $\Lambda'$  has successive minima  $s'_1 \leq \dots \leq s'_r$  and that  $\Lambda$  has successive minima  $s_1 \leq \dots \leq s_r$ . Then we have*

$$s'_i \geq s_i, \quad (1 \leq i \leq r),$$

*and  $\det \Lambda' \gg \det \Lambda$ .*

*Proof.* To prove the first part of the lemma we suppose for a contradiction that

$$s'_1 \leq \dots \leq s'_i < s_i,$$

for some  $1 \leq i \leq r$ . Now let  $\mathbf{m}'_1, \dots, \mathbf{m}'_r \in \Lambda'$  be the basis vectors described in Lemma 1, and analogously for  $\mathbf{m}_1, \dots, \mathbf{m}_r \in \Lambda$ . Then it follows from the construction of the successive minima that the  $i$  linearly independent vectors  $\mathbf{m}'_1, \dots, \mathbf{m}'_i$  must belong to the  $\mathbb{Z}$ -linear span of  $\mathbf{m}_1, \dots, \mathbf{m}_{i-1}$ . This is clearly impossible, which thereby establishes that  $s'_i \geq s_i$  for  $1 \leq i \leq r$ .

Finally we note that the second part of the lemma follows from the first part and (2.2).  $\square$

Next we recall a result of the second author [9, Proof of Theorem 4]. We shall say that a non-zero form defined over  $\mathbb{Z}$  is primitive if the highest common factor of its coefficients is 1.

**Lemma 3.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a primitive form of degree  $d \geq 2$ . Then either*

$$\|F\| \ll B^\theta, \quad \theta = d \binom{n-1+d}{n-1},$$

*or else there exists a form  $G \in \mathbb{Z}[x_1, \dots, x_n]$  of degree  $d$ , not proportional to  $F$ , such that every  $\mathbf{x} \in S(F; B, \mathbb{Z}^n)$  also satisfies the equation  $G(\mathbf{x}) = 0$ .*

We shall also need a number of basic results about the counting function for points on varieties. Let  $N \geq 3$  and let  $Y \subset \mathbb{P}^{N-1}$  be a variety. Recall that for any  $\mathbf{x} \in \mathbb{Z}^N$  we write  $[\mathbf{x}]$  for the corresponding point in  $\mathbb{P}^{N-1}(\mathbb{Q})$ . We define the height of any rational point  $x \in \mathbb{P}^{N-1}(\mathbb{Q})$  to be

$$H(x) = \max_{1 \leq i \leq N} |x_i|,$$

provided that  $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N$  satisfies  $x = [\mathbf{x}]$ . We shall write  $Y(\mathbb{Q})$  for  $Y \cap \mathbb{P}^{N-1}(\mathbb{Q})$ , and

$$N_Y(B) = \#\{x \in Y(\mathbb{Q}) : H(x) \leq B\}.$$

Thus in the case of Corollary 1, for example, we have  $N_X(B) = \frac{1}{2}N(F; B)$ , since  $\mathbf{x}$  and  $-\mathbf{x}$  represent the same point in projective space.

Our next result estimates  $N_M(B)$ , for an  $m$ -dimensional linear space  $M$  in  $\mathbb{P}^{N-1}$ .

**Lemma 4.** *Let  $M \subseteq \mathbb{P}^{N-1}$  be an  $m$ -dimensional linear space defined over  $\mathbb{Q}$ . Write*

$$\mathbf{M} = \{\mathbf{x} \in \mathbb{Z}^N : [\mathbf{x}] \in M\} \cup \{\mathbf{0}\},$$

*so that  $\mathbf{M}$  is a lattice of dimension  $m+1$ , and let  $s_1, \dots, s_{m+1}$  be the successive minima of  $\mathbf{M}$ . Then*

$$N_M(B) \ll \prod_{j=1}^{m+1} \left(1 + \frac{B}{s_j}\right).$$

*Moreover if  $m = 1$  then*

$$N_M(B) \ll 1 + \frac{B^2}{\det \mathbf{M}}.$$

*Proof.* Choose a basis  $\mathbf{m}_1, \dots, \mathbf{m}_{m+1}$  for  $\mathbf{M}$  as in Lemma 1. Then if  $\mathbf{x} \in \mathbf{M}$  with  $\max_i |x_i| \leq B$ , we will have

$$\mathbf{x} = \sum_{j=1}^{m+1} \lambda_j \mathbf{m}_j$$

with  $\lambda_j \ll B/s_j$ . Thus the number of such points is

$$\ll \prod_{j=1}^{m+1} \left(1 + \frac{B}{s_j}\right),$$

as required.

When  $m = 1$  we see that

$$\mathbf{x} = \lambda_1 \mathbf{m}_1 + \lambda_2 \mathbf{m}_2$$

with  $\lambda_j \ll B/s_j$ . When  $s_2 \gg B$  this implies that  $\lambda_2 = 0$ , and since  $N_M(B)$  counts only primitive vectors  $\mathbf{x}$  we must have  $\mathbf{x} = \pm \mathbf{m}_1$ . Thus  $N_M(B) \ll 1$  when  $s_2 \gg B$ . On the other hand, when  $s_1 \leq s_2 \ll B$  we find that

$$N_M(B) \ll \left(1 + \frac{B}{s_1}\right) \left(1 + \frac{B}{s_2}\right) \ll 1 + \frac{B^2}{s_1 s_2} \ll 1 + \frac{B^2}{\det \mathbf{M}},$$

by (2.2). This suffices for the lemma.  $\square$

We shall also need the following estimates for points on curves and surfaces.

**Lemma 5.** *Let  $C$  be an irreducible curve in  $\mathbb{P}^{N-1}$  of degree  $D \geq 2$ . Then*

$$N_C(B) \ll_{D,N} B^{2/D+\varepsilon} \ll_{D,N} B^{1+\varepsilon}.$$

*Similarly, if  $S \subset \mathbb{P}^{N-1}$  is any irreducible surface of degree  $D \geq 2$ , we have*

$$N_S(B) \ll_{D,N} B^{2+\varepsilon}.$$

The first assertion may be found in Broberg [1, Corollary 1], for example, while the second is in the first author's work [3, Lemma 1].

We end this section by establishing Theorem 1, for which we let  $\hat{Y} \subset \mathbb{A}^N$  be the affine cone above  $Y$ . Then  $\hat{Y}$  is an irreducible affine variety of degree  $D$  and dimension  $m+1$ . Now for any irreducible affine variety  $T \subset \mathbb{A}^\nu$  of degree  $\delta$  and dimension  $\mu$ , we define the quantity

$$M_T(B) = \#\{\mathbf{t} \in \mathbb{Z}^\nu : \mathbf{t} \in T, \max_i |t_i| \leq B\},$$

and proceed by proving that

$$M_T(B) = O_{\delta,\nu}(B^\mu). \quad (2.3)$$

This will clearly suffice to establish Theorem 1, since

$$N_Y(B) \leq M_{\hat{Y}}(B).$$

We shall establish (2.3) by induction on  $\mu$ . Since an irreducible variety of dimension zero contains just one point, the estimate is trivial whenever  $\mu = 0$ . Assume that  $\mu \geq 1$ . Since  $T$  is irreducible we may find an index  $1 \leq a \leq \nu$  such that  $T$  intersects the hyperplane  $t_a = \alpha$  properly, for any  $\alpha \in \mathbb{C}$ . Let  $H_\alpha$  denote this hyperplane. We thereby obtain the upper bound

$$M_T(B) \leq \sum_{|\alpha| \leq B} M_{T \cap H_\alpha}(B).$$

Since  $T \cap H_\alpha$  has dimension at most  $\mu - 1$  for every  $\alpha$ , and decomposes into at most  $D$  irreducible exponents, an application of the induction hypothesis implies that  $M_{T \cap H_\alpha}(B) = O_{\delta,\nu}(B^{\mu-1})$ . This suffices to complete the proof of (2.3), and so completes the proof of Theorem 1.

### 3 Deduction of Theorem 2 and Corollary 1

In this section we use Theorem 4 to deduce Theorem 2 and Corollary 1. Recall the definition (1.2) of  $V$  and let

$$b_i = \lfloor V/B_i \rfloor,$$

where  $\lfloor \alpha \rfloor$  denotes the integer part of any  $\alpha \in \mathbb{R}$ . We define the lattice

$$\Gamma(\mathbf{B}; n) = \{\mathbf{x} \in \mathbb{Z}^n : b_i \mid x_i, (1 \leq i \leq n)\}.$$



In particular  $\Gamma(\mathbf{B}; n)$  has rank  $n$ , successive minima  $b_1, \dots, b_n$  (though not necessarily in that order) and determinant

$$\det \Gamma(\mathbf{B}; n) = \prod_{i=1}^n \lfloor V/B_i \rfloor \gg V^{n-1}. \quad (3.1)$$

With these definitions in mind, we see that under the linear transformation  $\text{Diag}(b_1, \dots, b_n)$ , the image of the region  $|x_i| \leq B_i$  lies inside a cube with sides at most  $2V$ . This therefore establishes the useful inequality

$$N(F; \mathbf{B}) \ll \#S(F; V, \Gamma(\mathbf{B}, n)),$$

allowing us to move from considering points contained in a lopsided region to points contained in a suitable cube. In order to apply Theorem 4, we observe that the largest successive minimum  $s_n$  of  $\Gamma(\mathbf{B}; n)$  satisfies  $s_n \leq V$ . Indeed, if we assume that  $B_1 \geq \dots \geq B_n$  say, then  $s_i = b_i$  for  $1 \leq i \leq n$ . Now if  $\mathbf{x}$  lies on some linear space  $M \subseteq X$  of dimension at least 1, then it trivially lies on some projective line contained in  $X$ , and so is not to be counted by  $N_1(F; \mathbf{B})$ . Taking  $B = V$  and  $\Lambda = \Gamma(\mathbf{B}; n)$  in the statement of Theorem 4, we therefore deduce the first assertion of Theorem 2, via (3.1).

In order to deduce the second part of Theorem 2 we employ Lemma 3. Since we may assume that  $F$  is primitive, we deduce that either  $\log \|F\| \ll \log B$ , in which case we are done, or else that there exists a non-proportional form  $G$ , of degree  $d$ , such that every  $\mathbf{x} \in S(F; B, \mathbb{Z}^n)$  lies on the intersection  $F = G = 0$ . This is a union of  $O(1)$  irreducible varieties of codimension 2 in  $\mathbb{P}^{n-1}$ , and degree at most  $d^2$ . In this case the required estimate therefore follows from Theorem 1. This completes the proof of Theorem 2.

To deduce Corollary 1 we observe that it suffices, after Theorem 2, to control the number of points lying on lines contained in  $X$ . For any positive integer  $m \leq n-1$  we let  $\mathbb{G}(m, n-1)$  denote the variety parameterising  $m$ -dimensional linear subspaces of  $\mathbb{P}^{n-1}$ . It is well-known that  $\mathbb{G}(m, n-1)$  can be embedded in  $\mathbb{P}^{N-1}$  via the Plücker embedding, where  $N = \binom{n}{m+1}$ . Furthermore for any irreducible hypersurface  $Y$  in  $\mathbb{P}^{n-1}$  we let

$$F_m(Y) = \{M \in \mathbb{G}(m, n-1) : M \subset Y\}$$

denote the corresponding Fano variety of  $m$ -dimensional linear spaces contained in  $Y$ . Here we follow a common abuse of notation and identify linear spaces  $M \subset \mathbb{P}^{n-1}$  of dimension  $m$  with the corresponding point  $M \in \mathbb{G}(m, n-1)$ . We now record the following basic result, which allows us to control the degrees of several varieties relating to  $F_m(Y)$ , and which will also be used in §§9,10 below.

**Lemma 6.** *Let  $Y \subset \mathbb{P}^{n-1}$  be an irreducible hypersurface of degree  $D$  and let  $y \in Y$ . Then the three varieties*

$$F_m(Y), \quad \bigcup_{M \in F_m(Y)} M, \quad \{M \in F_m(Y) : y \in M\},$$

*all have degrees which can be bounded in terms of  $D$  and  $n$  alone.*

*Proof.* That the degree of  $F_m(Y)$  may be bounded in terms of  $D$  and  $n$  alone follows by using the defining equation for  $Y$  to write down the explicit equations

for  $F_m(Y)$ . To see the second part it will suffice to establish that the degree of  $\bigcup_{M \in \Psi} M$  does not exceed the degree of  $\Psi$ , for any irreducible component  $\Psi \subseteq F_m(Y)$ . But this follows from a straightforward modification to [6, Example 19.11], much along the lines of [2, Lemma 2.3.2]. Indeed this latter result establishes precisely this inequality in the case  $m = \dim \Psi = 1$ .

Finally we assume without loss of generality that  $y = [1, 0, \dots, 0]$ , and let  $p_{i_1 \dots i_{m+1}}$  be the Plücker coordinates of  $\mathbb{G}(m, n-1)$ , for  $1 \leq i_1 < \dots < i_{m+1} \leq n$ . Then it follows that the  $m$ -dimensional linear spaces in  $\mathbb{P}^{n-1}$  which pass through  $y$  are parameterised by the hyperplane

$$\Lambda : p_{i_1 \dots i_{m+1}} = 0, \quad (1 < i_1 < \dots < i_{m+1} \leq n),$$

in  $\mathbb{G}(m, n-1)$ . Hence

$$\{M \in F_m(Y) : y \in M\} = \Lambda \cap F_m(Y),$$

which clearly suffices to complete the proof of the lemma.  $\square$

We can now complete the proof of Corollary 1. By hypothesis we may assume that  $\bigcup_{L \in F_1(X)} L$  is a proper subvariety of  $X$ . Moreover Lemma 6 ensures that it is a union of  $O(1)$  irreducible varieties, each of degree  $O(1)$ . The result therefore follows from Theorem 1.

## 4 Conics

In this section we establish Theorem 6. We may assume at the outset that  $Q$  is primitive, since we can always remove any factor common to the coefficients in the equation  $Q = 0$ . Let  $\Delta$  be the discriminant of  $Q$  and fix a choice of  $r \in \mathbb{N}$ . (In fact  $r = 109$  will suffice.) By Bertrand's postulate we can choose primes  $p_1, \dots, p_r$ , with

$$cV^{1/3} \leq p_1 < \dots < p_r \ll_r V^{1/3}, \quad (4.1)$$

where  $c$  is an absolute constant to be chosen in due course. Now either there exists  $1 \leq i \leq r$  for which  $p_i \nmid \Delta$ , or else

$$|\Delta| \geq \prod_{i=1}^r p_i \gg V^{r/3}.$$

We begin by disposing of the latter case. Since  $|\Delta| \leq 6||Q||^3$  it will follow that

$$||Q|| \gg V^{r/9} \gg (\max B_i)^{r/9}.$$

We now apply Lemma 3 with  $d = 2$  and  $n = 3$ , so that  $\theta = 12$ . This shows that if  $r > 108$  then there exists a ternary quadratic form  $R$ , not proportional to  $Q$ , such that every  $\mathbf{x}$  counted by  $N(Q; \mathbf{B})$  also satisfies the equation  $R(\mathbf{x}) = 0$ . But then Bézout's theorem reveals that  $N(Q; \mathbf{B}) \leq 4$ , which is satisfactory.

We may now concentrate on the case in which there is a prime  $p$  in the range  $cV^{1/3} \leq p \ll V^{1/3}$  given by (4.1), with the property that  $p \nmid \Delta$ . Our argument now depends on the following lemma. We state the result in more generality than is needed for the proof of Theorem 6, so that it may be applied later in the treatment of Theorem 4.

**Lemma 7.** Let  $H(x_1, \dots, x_m) \in \mathbb{Z}[x_1, \dots, x_m]$  be a form of degree  $d$ . Let  $q$  be a power of a prime, and let  $\mathbf{x} \in \mathbb{Z}^m$  be a vector for which  $q \mid H(\mathbf{x})$  and such that

$$\text{h.c.f}\left(q, \frac{\partial H(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial H(\mathbf{x})}{\partial x_m}\right) = 1. \quad (4.2)$$

Then there is at least one vector  $\mathbf{x}^{(1)} \in \mathbb{Z}^m$  satisfying both (4.2) and  $q^2 \mid H(\mathbf{x}^{(1)})$ , and for which  $\mathbf{x}^{(1)} \equiv \mathbf{x} \pmod{q}$ . Write

$$\mathbf{M} = \mathbf{M}(\mathbf{x}, q, H) = \{\mathbf{w} \in \mathbb{Z}^m : \mathbf{w} \equiv \rho \mathbf{x} \pmod{q} \text{ some } \rho \in \mathbb{Z}, q^2 \mid \mathbf{w} \cdot \nabla H(\mathbf{x}^{(1)})\}.$$

Then  $\mathbf{M}$  is independent of the choice of  $\mathbf{x}^{(1)}$ . Moreover  $\mathbf{M}$  is a lattice of dimension  $m$  and determinant  $\det \mathbf{M} = q^m$ . Finally, if  $\mathbf{t} \in \mathbb{Z}^m$  satisfies  $H(\mathbf{t}) = 0$ , and if there exists  $\lambda \in \mathbb{Z}$  for which  $\mathbf{t} \equiv \lambda \mathbf{x} \pmod{q}$ , then  $\mathbf{t} \in \mathbf{M}$ .

*Proof.* The existence of a suitable  $\mathbf{x}^{(1)}$  is an immediate consequence of Hensel's Lemma. Specifically, on writing  $\mathbf{x}^{(1)} = \mathbf{x} + q\mathbf{y}^{(1)}$ , we find that  $q^2 \mid H(\mathbf{x}^{(1)})$  if and only if  $\mathbf{y}^{(1)} \cdot \nabla H(\mathbf{x}) \equiv -q^{-1}H(\mathbf{x}) \pmod{q}$ , and this is always solvable for  $\mathbf{y}^{(1)}$ , by (4.2). Now suppose that  $\mathbf{x}^{(2)} = \mathbf{x} + q\mathbf{y}^{(2)}$ , with  $\mathbf{y}^{(2)} \cdot \nabla H(\mathbf{x}) \equiv -q^{-1}H(\mathbf{x}) \pmod{q}$ , and moreover that  $\mathbf{w} \equiv \rho \mathbf{x} \pmod{q}$ . To show that  $\mathbf{M}$  is independent of the choice of  $\mathbf{x}^{(1)}$  it will suffice to demonstrate that  $q^2 \mid \mathbf{w} \cdot \nabla H(\mathbf{x}^{(1)})$  if and only if  $q^2 \mid \mathbf{w} \cdot \nabla H(\mathbf{x}^{(2)})$ . However

$$\begin{aligned} \mathbf{w} \cdot \nabla H(\mathbf{x}^{(i)}) &= \mathbf{w} \cdot \nabla H(\mathbf{x} + q\mathbf{y}^{(i)}) \\ &\equiv \mathbf{w} \cdot \nabla H(\mathbf{x}) + q \sum_{j,k=1}^n w_j y_k^{(i)} \frac{\partial^2 H}{\partial x_j \partial x_k}(\mathbf{x}) \pmod{q^2} \\ &\equiv \mathbf{w} \cdot \nabla H(\mathbf{x}) + q \sum_{j,k=1}^n \rho x_j y_k^{(i)} \frac{\partial^2 H}{\partial x_j \partial x_k}(\mathbf{x}) \pmod{q^2} \\ &\equiv \mathbf{w} \cdot \nabla H(\mathbf{x}) + q(d-1)\rho \mathbf{y}^{(i)} \cdot \nabla H(\mathbf{x}) \pmod{q^2} \\ &\equiv \mathbf{w} \cdot \nabla H(\mathbf{x}) - (d-1)\rho H(\mathbf{x}) \pmod{q^2} \end{aligned}$$

for  $i = 1, 2$ . It follows that  $q^2 \mid \mathbf{w} \cdot \nabla H(\mathbf{x}^{(1)})$  if and only if  $q^2 \mid \mathbf{w} \cdot \nabla H(\mathbf{x}^{(2)})$ , as required.

It is trivial that  $\mathbf{M}$  is a lattice, since it is clearly closed under addition. Moreover if  $\mathbf{y} \in \mathbb{Z}^m$  then  $q^2 \mathbf{y} \in \mathbf{M}$  (taking  $\rho = 0$ ), whence  $\mathbf{M}$  must have dimension  $m$ . To compute  $\det \mathbf{M}$  we observe that if we put  $\mathbf{w} = \rho \mathbf{x}^{(1)} + q\mathbf{z}$ , then we have

$$\begin{aligned} \mathbf{w} \cdot \nabla H(\mathbf{x}^{(1)}) &= \rho \mathbf{x}^{(1)} \cdot \nabla H(\mathbf{x}^{(1)}) + q\mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \\ &= \rho dH(\mathbf{x}^{(1)}) + q\mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \\ &\equiv q\mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \pmod{q^2}, \end{aligned}$$

whence

$$\mathbf{M} = \{\mathbf{w} \in \mathbb{Z}^m : \mathbf{w} = \rho \mathbf{x}^{(1)} + q\mathbf{z} \text{ some } \rho \in \mathbb{Z}, \mathbf{z} \in \mathbb{Z}^m, q \mid \mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)})\}.$$

We now observe that we have an inclusion of lattices  $q^2 \mathbb{Z}^m \subseteq \mathbf{M} \subseteq \mathbb{Z}^m$ . Hence in order to calculate the determinant  $\det \mathbf{M}$ , which is equal to the index of  $\mathbf{M}$

in  $\mathbb{Z}^m$  as an additive subgroup, it will suffice to calculate the index  $[\mathbf{M} : q^2\mathbb{Z}^m]$ . Indeed we then have

$$\det \mathbf{M} = [\mathbb{Z}^m : \mathbf{M}] = \frac{[\mathbb{Z}^m : q^2\mathbb{Z}^m]}{[\mathbf{M} : q^2\mathbb{Z}^m]} = \frac{q^{2m}}{[\mathbf{M} : q^2\mathbb{Z}^m]}. \quad (4.3)$$

We begin by considering the cosets of  $\mathbf{M}$  modulo  $q^2\mathbb{Z}^m$ . In view of the coprimality constraint (4.2) there are  $q^{m-1}$  possible values for  $\mathbf{z}$  modulo  $q$  satisfying  $q \mid \mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)})$ . Moreover there are  $q^2$  possible values for  $\rho$  modulo  $q^2$ . Finally, each value of  $\mathbf{w}$  may be decomposed as  $\mathbf{w} \equiv \rho \mathbf{x}^{(1)} + q\mathbf{z} \pmod{q^2}$  in  $q$  ways. It follows that  $\mathbf{M}(\mathbf{x}, q, H)$  has exactly  $q^m$  cosets modulo  $q^2\mathbb{Z}^m$ , and so (4.3) implies that  $\det \mathbf{M} = q^{2m} q^{-m} = q^m$ , as required.

For the final part of the lemma we note that if  $\mathbf{t} \equiv \lambda \mathbf{x} \pmod{q}$ , then  $\mathbf{t} \equiv \lambda \mathbf{x}^{(1)} \pmod{q}$ . Thus there is a vector  $\mathbf{z} \in \mathbb{Z}^m$  such that  $\mathbf{t} = \lambda \mathbf{x}^{(1)} + q\mathbf{z}$ . It follows that

$$0 = H(\mathbf{t}) \equiv \lambda^d H(\mathbf{x}^{(1)}) + q\lambda^{d-1} \mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \equiv q\lambda^{d-1} \mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \pmod{q^2}.$$

Moreover we must have  $\text{h.c.f.}(\lambda, q) = 1$  since  $\mathbf{t}$  is primitive and  $\mathbf{t} \equiv \lambda \mathbf{x} \pmod{q}$ . This allows us to conclude that  $q \mid \mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)})$ . We must then have

$$\begin{aligned} \mathbf{t} \cdot \nabla H(\mathbf{x}^{(1)}) &= \lambda \mathbf{x}^{(1)} \cdot \nabla H(\mathbf{x}^{(1)}) + q\mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \\ &= \lambda d H(\mathbf{x}^{(1)}) + q\mathbf{z} \cdot \nabla H(\mathbf{x}^{(1)}) \\ &\equiv 0 \pmod{q^2}, \end{aligned}$$

so that  $\mathbf{t} \in \mathbf{M}$ . This completes the proof of the lemma.  $\square$

We can now complete the proof of Theorem 6. We have shown that we may assume there is a prime  $p \nmid \Delta$  in the range  $cV^{1/3} \leq p \ll V^{1/3}$ . The projective variety  $Q(\mathbf{x}) = 0$  has exactly  $p$  points over  $\mathbb{F}_p$ , and we aim to show that there are at most 2 points counted by  $N(Q; \mathbf{B})$  lying above each one. This will suffice for the theorem. We therefore fix a vector  $\mathbf{x} \in Z^3$  with  $Q(\mathbf{x}) \equiv 0 \pmod{p}$ , and note that (4.2) holds for  $\mathbf{x}$  with  $q = p$  since  $p \nmid \Delta$ . Next we count vectors  $\mathbf{w} \in Z^3$  satisfying  $Q(\mathbf{w}) = 0$ ,

$$|w_1| \leq B_1, \quad |w_2| \leq B_2, \quad |w_3| \leq B_3. \quad (4.4)$$

and such that there exists  $\rho \in \mathbb{Z}$  with  $\mathbf{w} \equiv \rho \mathbf{x} \pmod{p}$ . According to Lemma 7 we will have  $\mathbf{w} \in \mathbf{M} = \mathbf{M}(\mathbf{x}, p, Q)$ . We now consider the map

$$\phi(y_1, y_2, y_3) = (B_2 B_3 y_1, B_1 B_3 y_2, B_1 B_2 y_3).$$

Then  $\phi(\mathbf{M})$  has determinant  $V^2 \det \mathbf{M} = V^2 p^3$ . Let  $s_1 \leq s_2 \leq s_3$  be the successive minima of  $\phi(\mathbf{M})$ , and let  $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$  be the corresponding basis vectors described in Lemma 1. It then follows from (2.2) that  $s_3 \geq V^{2/3} p \gg V$ . On the other hand, if (4.4) holds, then  $\phi(\mathbf{w})$  has Euclidean length  $\ll V$ . Hence Lemma 1 implies that, if we choose the absolute constant  $c$  in (4.1) to be sufficiently large, the vector  $\phi(\mathbf{w})$  must lie in the 2-dimensional lattice spanned by  $\mathbf{m}_1$  and  $\mathbf{m}_2$ . Thus, given  $\mathbf{x}$ , all corresponding points  $\mathbf{w}$  must lie not only on the conic  $Q(\mathbf{w}) = 0$  but also on a certain line. There are therefore at most 2 such points  $\mathbf{w} \in Z^3$ . This completes the proof of Theorem 6.

## 5 Proof of Theorem 4 — variable moduli

The purpose of this section is to examine how the ideas of the previous section may be adapted to hypersurfaces of dimension 2 or more. It is clear that §4 depends crucially on condition (4.2). For any prime  $p$ , any lattice  $\Lambda \subseteq \mathbb{Z}^n$  of dimension  $m \geq 2$ , and any form  $G \in \mathbb{Z}[x_1, \dots, x_n]$  of degree  $d$ , we therefore introduce the set

$$S(G; B, \Lambda, p) = \{\mathbf{x} \in S(G; B, \Lambda) : p \nmid \nabla G(\mathbf{x})\}.$$

We could then mimic the argument from §4 to show that if  $p \gg B$  then the points in  $S(G; B, \Lambda, p)$  lie in  $O(B^{m-2})$  sublattices  $\Lambda'$  of  $\Lambda$ , each of dimension  $m - 1$ . In §4, where  $G = Q$  is a non-singular ternary quadratic form and  $\Lambda = \mathbb{Z}^3$ , each sublattice  $\Lambda'$  corresponds to a line, and each line contains at most 2 points on the conic  $Q = 0$ . When  $m = 4$ , for example, the situation is more complicated. Each sublattice  $\Lambda'$  corresponds to a plane in  $\mathbb{P}^3$ , so that the corresponding points of  $G = 0$  typically lie on a curve. In order to estimate the number of such points we would want to know the size of  $\det \Lambda'$ , and in order to get a good estimate we would want  $\det \Lambda'$  to be large. Since it is clear that we produce different lattices  $\Lambda'$  for different points  $\mathbf{x} \pmod{p}$  we must anticipate difficulties in trying to make  $\det \Lambda'$  large for all  $\mathbf{x} \pmod{p}$ . We shall overcome these problems by using moduli which are powers  $p^k$  of  $p$ , and allowing different values of  $k$  for different vectors  $\mathbf{x}$ .

It is now time to state the result we shall use.

**Lemma 8.** *Let  $\Lambda \subseteq \mathbb{Z}^n$  be a lattice of dimension  $m \geq 2$ , with largest successive minimum at most  $B$ . Let  $p$  be a prime with  $p \nmid \det \Lambda$ . Then there is an integer  $I \ll p^{m(m-2)} (B^m / \det \Lambda)^{(m-2)/m}$  and lattices  $\Lambda_1, \dots, \Lambda_I \subset \Lambda$  of dimension  $m - 1$ , such that*

$$S(G; B, \Lambda, p) \subseteq \bigcup_{i=1}^I \Lambda_i.$$

For any  $i$ , the successive minima of  $\Lambda_i$  are all  $O(p^m B)$ . Moreover for any  $\alpha \geq 1$  we have

$$\#\{i : \det \Lambda_i \leq \alpha \det \Lambda\} \ll p^{m-2} (\alpha B)^{(m-2)/m}.$$

*Proof.* Let  $\mathbf{e}_1, \dots, \mathbf{e}_m$  be a basis for  $\Lambda$  and set

$$H(y_1, \dots, y_m) = G(y_1 \mathbf{e}_1 + \dots + y_m \mathbf{e}_m).$$

Let  $\phi : \Lambda \rightarrow \mathbb{Z}^m$  be the map

$$\phi : \mathbf{x} = y_1 \mathbf{e}_1 + \dots + y_m \mathbf{e}_m \mapsto \mathbf{y}.$$

Then

$$\frac{\partial H}{\partial y_j} = \mathbf{e}_j \cdot \nabla G(\mathbf{x}),$$

and since  $p \nmid \det \Lambda$  we see that  $p \nmid \nabla G(\mathbf{x})$  implies  $p \nmid \nabla H(\mathbf{y})$ .

We proceed by defining a finite tree for each non-singular projective point  $\mathbf{y}$  on the mod  $p$  reduction of  $H = 0$ . There will be  $O(p^{m-2})$  such points, uniformly in  $H$ . We define the “depth” of the root node to be 1, and if a node  $N'$  is an immediate successor to a node  $N$  we define the depth of  $N'$  to be 1 more than

the depth of  $N$ . This terminology may differ slightly from that employed by other authors.

We shall define an equivalence relation on the set

$$\{\mathbf{v} \in (\mathbb{Z}/p^k\mathbb{Z})^m : \text{h.c.f.}(v_1, \dots, v_m, p) = 1\},$$

by taking  $\mathbf{v} \sim_k \mathbf{v}'$  if and only if  $\mathbf{v} = \lambda \mathbf{v}'$  for some unit  $\lambda$  of  $\mathbb{Z}/p^k\mathbb{Z}$ . The tree we construct will then have each node of depth  $k$  labelled by an equivalence class  $[\mathbf{v}]_k$ , such that, if  $\mathbf{v} = \mathbf{w} + p^k \mathbb{Z}^m$  then  $p^k | H(\mathbf{w})$  and  $p \nmid \nabla H(\mathbf{w})$ . The root node is thus labelled by the equivalence class  $[\mathbf{y}]_1$ , where  $\mathbf{y}$  is the point mentioned above.

To a node with label  $[\mathbf{v}]_k$ , where  $\mathbf{v} = \mathbf{w} + p^k \mathbb{Z}^m$ , we associate the lattice  $\mathbf{N} = \phi^{-1}\mathbf{M}(\mathbf{w}, p^k, H)$ , in the notation of Lemma 7. Clearly the lattice will be independent of the choice of  $\mathbf{w}$ . We then declare the node to be a leaf if the largest successive minimum  $s_m$  of this lattice satisfies  $s_m > cB$ . Here  $c$  is an absolute constant whose value will become clear in due course.

We must now compute the determinant of  $\mathbf{N}$ . If we take  $\mathbf{y}_1, \dots, \mathbf{y}_m \in \mathbb{Z}^m$  to be column vectors forming a basis of  $\mathbf{M}(\mathbf{w}, p^k, H)$ , then  $\mathbf{N}$  will have a basis  $\mathbf{E}\mathbf{y}_1, \dots, \mathbf{E}\mathbf{y}_m$ , where  $\mathbf{E}$  is the  $n \times m$  matrix with columns  $\mathbf{e}_1, \dots, \mathbf{e}_m$ . Then (2.1) implies that

$$(\det \mathbf{N})^2 = \det(\mathbf{y}_i^T \mathbf{E}^T \mathbf{E} \mathbf{y}_j)$$

and

$$(\det \Lambda)^2 = \det(\mathbf{E}^T \mathbf{E}).$$

Since  $\mathbf{E}^T \mathbf{E}$  is a real positive symmetric matrix, we may write  $\mathbf{E}^T \mathbf{E} = \mathbf{F}^T \mathbf{F}$  for some real  $m \times m$  matrix  $\mathbf{F}$ . Thus if  $\mathbf{Y}$  is the  $m \times m$  matrix with columns  $\mathbf{y}_1, \dots, \mathbf{y}_m$  we will have

$$\begin{aligned} (\det \mathbf{N})^2 &= \det(\mathbf{y}_i^T \mathbf{E}^T \mathbf{E} \mathbf{y}_j) \\ &= \det(\mathbf{y}_i^T \mathbf{F}^T \mathbf{F} \mathbf{y}_j) \\ &= \det(\mathbf{Y}^T \mathbf{F}^T \mathbf{F} \mathbf{Y}) \\ &= (\det \mathbf{Y})^2 (\det \mathbf{F})^2, \end{aligned}$$

and

$$(\det \Lambda)^2 = \det(\mathbf{E}^T \mathbf{E}) = \det(\mathbf{F}^T \mathbf{F}) = (\det \mathbf{F})^2.$$

Finally,

$$\det \mathbf{M}(\mathbf{w}, p^k, H) = |\det(\mathbf{Y})|,$$

whence

$$\det \mathbf{N} = (\det \mathbf{M}(\mathbf{w}, p^k, H))(\det \Lambda) = p^{km} \det \Lambda, \quad (5.1)$$

by Lemma 7.

Thus the largest successive minimum of  $\mathbf{N}$  satisfies

$$s_m \geq (\det \mathbf{N})^{1/m} = p^k (\det \Lambda)^{1/m}, \quad (5.2)$$

by (2.2). For each node  $[\mathbf{v}]_k$  which is not a leaf, we take its immediate successors to be those nodes labelled by  $[\mathbf{u}]_{k+1}$ , for which  $\mathbf{u} = \mathbf{w} + p^{k+1} \mathbb{Z}^m \in (\mathbb{Z}/p^{k+1}\mathbb{Z})^m$  satisfies both  $\mathbf{v} = \mathbf{w} + p^k \mathbb{Z}^m$  and  $p^{k+1} | H(\mathbf{w})$ . In particular we have  $p \nmid \nabla H(\mathbf{w})$  since  $[\mathbf{v}]_k$  is a node. We claim that there will be precisely  $p^{m-2}$  such immediate successors. This is a simple application of Hensel's Lemma. Clearly it will

suffice to show that for any node  $[\mathbf{v}]_k$ , there are  $p^{m-1}$  vectors  $\mathbf{u} = \mathbf{w} + p^{k+1}\mathbb{Z}^m$  for which  $\mathbf{v} = \mathbf{w} + p^k\mathbb{Z}^m$  and  $p^{k+1} \mid H(\mathbf{w})$ . Suppose that  $\mathbf{v} = \mathbf{w} + p^k\mathbb{Z}^m$ , with  $p \nmid \nabla H(\mathbf{w})$  and  $H(\mathbf{w}) = p^k w$ , say. Then we count values of  $\mathbf{u} \in (\mathbb{Z}/p^{k+1}\mathbb{Z})^m$  for which  $\mathbf{u} \equiv \mathbf{w} \pmod{p^k}$  and  $p^{k+1} \mid H(\mathbf{u})$ . Writing  $\mathbf{u} = \mathbf{w} + p^k\mathbf{z}$ , we deduce that

$$H(\mathbf{u}) \equiv H(\mathbf{w}) + p^k \mathbf{z} \cdot \nabla H(\mathbf{w}) \pmod{p^{k+1}},$$

and so seek values of  $\mathbf{z}$  for which

$$w + \mathbf{z} \cdot \nabla H(\mathbf{w}) \equiv 0 \pmod{p}.$$

But this has exactly  $p^{m-1}$  solutions  $\mathbf{z} \pmod{p}$ , since  $p \nmid \nabla H(\mathbf{w})$ . Therefore there are  $p^{m-2}$  immediate successors above any given node which is not a leaf.

We will need to establish an upper bound for  $s_m$ . In fact we shall show that

$$s_m \ll p^m B. \quad (5.3)$$

Suppose first that the root node  $[\mathbf{y}]_1$  is a leaf, and let  $\mathbf{N} = \phi^{-1}\mathbf{M}(\mathbf{y}, p, H)$  be the corresponding lattice. Then upon observing that we have an inclusion of lattices  $p^2\Lambda \subseteq \mathbf{N}$ , it follows from Lemma 2 that the successive minima of  $\mathbf{N}$  are at most  $p^2$  times as large as those of  $\Lambda$ . Since the largest successive minimum of  $\Lambda$  is at most  $B$  by hypothesis, we therefore have  $s_m \leq p^2 B$ , which is satisfactory for (5.3). Whenever the root node is not a leaf we may consider the node  $[\mathbf{v}']_{k-1}$ , say, immediately preceding the leaf  $[\mathbf{v}]_k$  for some  $k \geq 2$ . Let the lattices corresponding to  $[\mathbf{v}]_k$  and  $[\mathbf{v}']_{k-1}$  be  $\mathbf{N}$  and  $\mathbf{N}'$  respectively, with successive minima  $s_1, \dots, s_m$  and  $s'_1, \dots, s'_m$ . We note that  $\mathbf{N} \subseteq \mathbf{N}'$ , as one sees directly from the definition in Lemma 7. But then it follows from Lemma 2 that

$$s_i \geq s'_i, \quad (1 \leq i \leq m).$$

Since  $[\mathbf{v}']_{k-1}$  is not a leaf we have  $s'_m \ll B$ . Therefore (2.2) yields

$$\begin{aligned} s_m &\ll \frac{\det \mathbf{N}}{s_1 \cdots s_{m-1}} \\ &\leq \frac{\det \mathbf{N}}{s'_1 \cdots s'_{m-1}} \\ &\leq \frac{\det \mathbf{N}}{\det \mathbf{N}' / s'_m}, \end{aligned}$$

whence an application of (5.1) completes the proof of (5.3).

We may conclude from (5.2) and (5.3) that

$$p^{k-m} \ll \left( \frac{B^m}{\det \Lambda} \right)^{1/m}.$$

In view of this we now see that the tree is finite, with total depth

$$k_0 \ll 1 + \frac{\log(B^m / \det \Lambda)}{\log p},$$

and has at most

$$p^{(m-2)(k_0-1)} \ll p^{(m-1)(m-2)} \left( \frac{B^m}{\det \Lambda} \right)^{(m-2)/m}$$

leaf nodes. The reader should also recall that we have  $O(p^{m-2})$  such trees to consider, giving  $O(p^{m(m-2)}(B^m/\det \Lambda)^{(m-2)/m})$  leaf nodes in total.

Now suppose that  $\mathbf{x} \in S(G; B, \Lambda, p)$ . Then  $\mathbf{y} = \phi(\mathbf{x})$  satisfies  $H(\mathbf{y}) = 0$  and  $p \nmid \nabla H(\mathbf{y})$ . The construction of the various trees is such that there is then a leaf node  $[\mathbf{v}]_k$  in one of the trees with the property that  $\mathbf{y} \equiv \lambda \mathbf{v} \pmod{p^k}$ , for some  $\lambda$ . It follows from Lemma 7 that  $\mathbf{y}$  belongs to the corresponding lattice  $\mathbf{M}(\mathbf{v}, p^k, H)$ , and hence that  $\mathbf{x}$  belongs to the corresponding lattice  $\mathbf{N}$ . As we have seen in Lemma 1, there exists a basis  $\mathbf{m}_1, \dots, \mathbf{m}_m$  of the lattice  $\mathbf{N}$  such that if one writes any  $\mathbf{x} \in \mathbf{N}$  as  $\mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{m}_i$ , then  $\lambda_i \ll |\mathbf{x}|/s_i$ . Thus if  $\max_i |x_i| \leq B$  and  $s_m > cB$  with  $c$  a suitably chosen constant, we must have  $\lambda_m = 0$ . Hence if  $\mathbf{x} \in S(G; B, \Lambda, p)$  corresponds to the leaf node  $[\mathbf{v}]_k$  then  $\mathbf{x}$  must belong to the lattice  $\Lambda_i$ , say, spanned by  $\mathbf{m}_1, \dots, \mathbf{m}_{m-1}$ . The number of such lattices  $\Lambda_i$  is the total number of leaf nodes, which is  $O(p^{m(m-2)}(B^m/\det \Lambda)^{(m-2)/m})$ .

We next consider the successive minima of  $\Lambda_i$ . Clearly these are no larger than the corresponding values of  $s_m$ , which by (5.3) are  $O(p^m B)$ . Indeed it is clear that any set of linearly independent vectors in  $\Lambda_i$  is also a set of linearly independent vectors in  $\mathbf{N}$ , whence the successive minima of  $\Lambda_i$  must be precisely  $s_1, \dots, s_{m-1}$ . According to (2.2) and (5.1) we therefore have

$$\det \Lambda_i \gg \prod_{j=1}^{m-1} s_j = s_m^{-1} \prod_{j=1}^m s_j \geq \frac{\det \mathbf{N}}{s_m} \gg p^{km-m} B^{-1} \det \Lambda.$$

Thus  $\det \Lambda_i \leq \alpha \det \Lambda$  implies that  $p^k \ll p(\alpha B)^{1/m}$ . Since each tree has at most  $p^{(m-2)(k-1)}$  leaf nodes of depth  $k$  it follows that there are at most  $O(p^{m-2}(\alpha B)^{(m-2)/m})$  possible values for  $i$ . This completes the proof of the lemma.  $\square$

For future use we also note that

$$\det \Lambda_i \leq \prod_{j=1}^{m-1} s_j = s_m^{-1} \prod_{j=1}^m s_j \ll \frac{\det \mathbf{N}}{s_m} \ll p^{km} B^{-1} \det \Lambda, \quad (5.4)$$

by a further application of (2.2), (5.1) and the fact that  $s_m \gg B$  by construction.

## 6 Proof of Theorem 4 — the induction

In this section we shall complete the proof of Theorem 4 using an induction argument. For this we shall fix  $n$ , and use induction on  $m$ . We take  $m = 2$  as the base for the induction. In this case  $S(F; B, \Lambda)$  consists of  $O(1)$  points  $\mathbf{t}$  with  $|\mathbf{t}| \ll B$ . Thus we can take  $J \ll 1$ , and each  $\mathbf{M}_j$  will be spanned by the corresponding  $\mathbf{t}$ . Moreover, each  $\mathbf{M}_j$  will have dimension 1 and successive minimum  $O(B)$ , so that (1.4) holds with  $\mathbf{w} = \mathbf{t}$  for any  $q_1$ . Choosing a sufficiently large prime  $q_1 \equiv 2 \pmod{p_0}$  will therefore suffice to ensure that (1.5) holds. Hence the required results hold for the base case  $m = 2$  and we may assume henceforth that  $m \geq 3$ .

Our first task is to show how Lemma 8 may be applied, and this is achieved via the following result.



**Lemma 9.** *Let  $\mathcal{P} = \log^2(\|F\|B)$ . Then there is an integer  $t \ll 1$  and forms  $G_1, \dots, G_t \in \mathbb{Z}[x_1, \dots, x_n]$ , with degrees at most  $d$ , not vanishing identically on  $\Lambda$ . Moreover*

$$\|G_j\| \ll \|F\|, \quad (1 \leq j \leq t),$$

and

$$S(F; B, \Lambda) \subseteq \bigcup_{1 \leq j \leq t} \bigcup_{\mathcal{P} \leq p \ll \mathcal{P}} S(G_j; B, \Lambda, p). \quad (6.1)$$

Here the primes  $p$  are restricted to satisfy the conditions  $p \equiv m \pmod{p_0}$  and  $p \nmid \det \Lambda$ .

*Proof.* We begin by defining

$$\mathcal{G} = \left\{ \left. \frac{\partial^{a_1+\dots+a_n} F}{\partial^{a_1} x_1 \dots \partial^{a_n} x_n} \right|_{\Lambda} \neq 0 : (a_1, \dots, a_n) \in \mathbb{Z}_{\geq 0}^n \right\}.$$

Thus  $\mathcal{G}$  is just the set of all partial derivatives of  $F$  which do not vanish identically on  $\Lambda$ . We then take the forms  $G_j$  to be the forms in the set  $\mathcal{G}$ . Thus it suffices to show (6.1), the remaining claims being obvious.

Let  $\mathbf{x} \in S(F; B, \Lambda)$  and consider the subset

$$\mathcal{G}(\mathbf{x}) = \{G \in \mathcal{G} : G(\mathbf{x}) = 0\}.$$

In particular we observe that  $\mathcal{G}(\mathbf{x})$  is non-empty since  $F \in \mathcal{G}$ . We take  $G = G_{\mathbf{x}}$  to be any form in the set  $\mathcal{G}(\mathbf{x})$  which has minimal degree, and proceed to show that  $\mathbf{x}$  is a non-singular point of  $G$ . To see this we first note that  $G(\mathbf{x}) = 0$ , since  $G \in \mathcal{G}(\mathbf{x})$ . Moreover, our choice of  $\mathcal{G}$  implies that  $G$  does not vanish identically on  $\Lambda$ . Hence  $G$  must have degree at least 1, since  $G(\mathbf{x}) = 0$ . But then it follows that the components of  $\nabla G$  cannot all vanish identically on  $\Lambda$ , and so there exists  $1 \leq i \leq n$  for which the partial derivative  $\partial G / \partial x_i$  does not vanish identically on  $\Lambda$ . Finally we deduce from the minimality of  $\deg G$  that

$$\frac{\partial G}{\partial x_i}(\mathbf{x}) \neq 0,$$

from which it follows that  $\nabla G(\mathbf{x}) \neq \mathbf{0}$  as claimed. It is perhaps instructive to remark at this point that if  $F$  is a non-singular form, as it was in §4, then we can take  $t = 1$  and  $G_1 = F$ . In fact our construction of auxiliary forms  $G_j$  is used purely in order to handle the contribution to  $S(F; B, \Lambda)$  from the singular locus of  $F = 0$ .

Now let  $G = G_{\mathbf{x}}$  be the form constructed above, with  $\nabla G(\mathbf{x}) \neq \mathbf{0}$ , and let  $p$  be the least prime number  $p \geq \mathcal{P}$  for which  $p \equiv m \pmod{p_0}$  and

$$p \nmid (\det \Lambda) \nabla G(\mathbf{x}).$$

We now observe that

$$\det \Lambda \leq \prod_{j=1}^m s_j \leq B^m,$$

by (2.2) and the hypotheses of Theorem 4, whence  $\log \det \Lambda \ll \log B$ . Recalling that  $G$  is non-singular at  $\mathbf{x}$ , a trivial modification to the proof of [9, Lemma 4] then shows that  $p$  exists, and that

$$\mathcal{P} \leq p \ll \mathcal{P}.$$

It follows that  $\mathbf{x}$  belongs to the corresponding set  $S(G; B, \Lambda, p)$ , and this suffices for the lemma.  $\square$

We are now ready to begin the proof of Theorem 4, using induction on  $m$ . According to Lemmas 8 and 9 we have

$$S(F; B, \Lambda) \subseteq \bigcup_{1 \leq j \leq t} \bigcup_{\mathcal{P} \leq p \ll \mathcal{P}} \bigcup_{i=1}^I \Lambda_i$$

for suitable lattices  $\Lambda_1, \dots, \Lambda_I \subset \Lambda$  of dimension  $m-1$ . Moreover the successive minima of these lattices are all

$$\leq c\mathcal{P}^m B = B_0, \quad (6.2)$$

say, for a suitable constant  $c$  depending only on  $d$  and  $n$ . Hence

$$S(F; B, \Lambda) \subseteq \bigcup_{\mathbf{M}} S(F; B_0, \mathbf{M}),$$

where  $\mathbf{M}$  runs over all the lattices  $\Lambda_i$  as the forms  $G_j$  and the primes  $p$  vary. The number of lattices  $\mathbf{M}$  with  $\det \mathbf{M} \leq \alpha \det \Lambda$  is

$$O(\mathcal{P}^{m-1}(\alpha B)^{(m-2)/m}),$$

since there are  $O(\mathcal{P})$  primes overall. Similarly, one finds that the total number of lattices  $\mathbf{M}$  is  $O(\mathcal{P}^{m(m-2)+1}(B^m / \det \Lambda)^{(m-2)/m})$ . It follows from (6.2) that

$$\det \mathbf{M} \ll \mathcal{P}^{m(m-1)} B^{m-1}. \quad (6.3)$$

We now focus on a particular lattice  $\mathbf{M}$ . If  $\mathbf{M} = \Lambda_i$  arises from  $\mathbf{M}(\mathbf{w}, p^k, H)$  in the proof of Lemma 8, then we take  $q_1 = p^k$ . In particular it follows that  $q_1$  is a power of a prime  $p_1 = p \equiv m \pmod{p_0}$ . Our construction gives

$$\Lambda_i \subset \mathbf{N} = \phi^{-1} \mathbf{M}(\mathbf{w}, p^k, H),$$

from which it is clear that if  $\mathbf{x} \in \Lambda_i$  then

$$\mathbf{x} \equiv \rho \phi^{-1}(\mathbf{w}) \pmod{q_1} \quad (6.4)$$

for some  $\rho \in \mathbb{Z}$ . If the form  $F$  vanishes on the lattice  $\mathbf{M}$  then the variety  $F = 0$  contains the corresponding linear space. The number of such lattices is satisfactory for (1.3), and (6.4) suffices for (1.4). Moreover, (5.4) suffices for (1.5), since  $h = 1$ .

We suppose from now on that  $F$  does not vanish identically on  $\mathbf{M}$ . We may therefore apply our induction hypothesis to  $S(F; B_0, \mathbf{M})$  and conclude that it is covered by

$$\begin{aligned} &\ll \left( \frac{B_0^{m-1}}{\det \mathbf{M}} \right)^{(m-3)/(m-1)} (\log \|F\| B)^{(m-3)\{2(m-1)^2 + (m-1) + 3\}/3} \\ &\ll \left( \frac{B^{m-1}}{\det \mathbf{M}} \right)^{(m-3)/(m-1)} \mathcal{P}^{m(m-3)} (\log \|F\| B)^{(m-3)\{2(m-1)^2 + (m-1) + 3\}/3} \end{aligned}$$

linear spaces, each contained in the variety  $F = 0$ . It follows that the total number of linear subvarieties arising from all those lattices  $\mathbf{M}$  for which

$$\frac{1}{2}\alpha \det \Lambda < \det \mathbf{M} \leq \alpha \det \Lambda$$

is

$$\begin{aligned} &\ll \mathcal{P}^{m-1}(\alpha B)^{(m-2)/m} \\ &\quad \times \left( \frac{B^{m-1}}{\alpha \det \Lambda} \right)^{(m-3)/(m-1)} \mathcal{P}^{m(m-3)} (\log \|F\|B)^{(2m^3-9m^2+13m-12)/3} \\ &\ll \frac{\alpha^{2/(m(m-1))} B^{(m^2-2m-2)/m}}{(\det \Lambda)^{(m-3)/(m-1)}} (\log \|F\|B)^{(2m^3-3m^2+m-18)/3}. \end{aligned}$$

In view of (6.3) we must sum this over dyadic ranges for

$$\alpha \ll \frac{B^{m-1}}{\det \Lambda} \mathcal{P}^{m(m-1)},$$

producing a total

$$\ll \left( \frac{B^m}{\det \Lambda} \right)^{(m-2)/m} (\log \|F\|B)^{(2m^3-3m^2+m-6)/3},$$

as required. This establishes the bound (1.3).

By our induction hypothesis, the successive minima of the lattices that arise will be

$$\ll B_0 (\log \|F\|B_0)^{(m-1)m}.$$

In view of our choice of  $B_0$  this is  $O(B(\log \|F\|B)^{m(m+1)})$ , as required.

According to our induction hypothesis  $S(F; B_0, \mathbf{M})$  is covered by lattices  $\mathbf{M}_j \subseteq \mathbf{M}$ . When  $\mathbf{M}_j$  has dimension  $(m-1) - h$  for  $h \geq 1$ , there are associated to it positive integers  $q'_1, \dots, q'_h$ , and a vector  $\mathbf{w}' \in \mathbb{Z}^n$  such that

$$\mathbf{M}_j \subseteq \{\mathbf{x} \in \mathbf{M} : \mathbf{x} \equiv \rho \mathbf{w}' \pmod{q'_1 \cdots q'_h} \text{ some } \rho \in \mathbb{Z}\}.$$

Each  $q'_i$  is a power of a prime  $p'_i \equiv (m-1) + 1 - i \pmod{p_0}$ . Now we have already noted in (6.4) that there exists an integer  $q_1$  which is a power of a prime  $p_1 \equiv m \pmod{p_0}$ , such that

$$\mathbf{x} \equiv \rho \phi^{-1}(\mathbf{w}) \pmod{q_1}$$

for some  $\rho \in \mathbb{Z}$ , whenever  $\mathbf{x} \in \mathbf{M}$ . For  $1 \leq i \leq h+1$  we define the integers

$$q_i = \begin{cases} q_1, & i = 1, \\ q'_{i-1}, & i > 1. \end{cases}$$

In particular it follows that each  $q_i$  is a power of a prime  $p_i \equiv m+1-i \pmod{p_0}$ . These congruence constraints ensure that  $q_1$  is coprime to  $q_2 \cdots q_{h+1}$ . Hence the Chinese Remainder Theorem implies that there is a vector  $\mathbf{w} \in \mathbb{Z}^n$  such that

$$\mathbf{M}_j \subseteq \{\mathbf{x} \in \mathbf{M} : \mathbf{x} \equiv \rho \mathbf{w} \pmod{q_1 q_2 \cdots q_{h+1}} \text{ some } \rho \in \mathbb{Z}\},$$

as required for (1.4).

It remains to consider the bound (1.5). However on applying (1.5) for  $M$  we find that

$$\det M_j \ll (\det M) \prod_{i=1}^h \frac{q_i^{t(m-1)-(i-1)}}{B_0} = (\det M) \prod_{i=1}^h \frac{q_i^{m-i}}{B_0}.$$

Since  $B_0 \gg B$ , the bound (5.4) yields

$$\det M_j \ll q_1^m B^{-1} (\det \Lambda) \prod_{i=2}^{h+1} \frac{q_i^{m-(i-1)}}{B} \ll (\det \Lambda) \prod_{i=1}^{h+1} \frac{q_i^{m-(i-1)}}{B},$$

as required. This completes the proof of Theorem 4.

## 7 Proof of Theorem 5

We begin by defining the height of a linear space. For positive integers  $m \leq n-1$ , we recall from §3 the variety  $\mathbb{G}(m, n-1)$  which parameterises  $m$ -dimensional linear subspaces of  $\mathbb{P}^{n-1}$ , and which may be embedded in projective space via the Plücker embedding. Whenever  $M \in \mathbb{G}(m, n-1)$  is defined over  $\mathbb{Q}$ , we define the height  $H(M)$  of  $M$  to be the height of its coordinates in  $\mathbb{G}(m, n-1)$ , under the Plücker embedding. Then for any  $M \in \mathbb{G}(m, n-1)$  which is defined over  $\mathbb{Q}$ , we let

$$M = \{\mathbf{x} \in \mathbb{Z}^n : [\mathbf{x}] \in M\} \cup \{\mathbf{0}\}$$

be the lattice associated to  $M$ . It follows from §2 that there exists a basis  $\mathbf{e}_1, \dots, \mathbf{e}_{m+1}$  of  $M$  such that  $|\mathbf{e}_1| \cdots |\mathbf{e}_{m+1}|$  has the same order of magnitude as  $\det M$ . In fact it can be shown (see Schmidt [11, Chapter I, Corollary 5I], for example) that

$$\det M = H(M). \quad (7.1)$$

The proof of Theorem 5 is now straightforward. Suppose that  $\Lambda = \mathbb{Z}^n$  in Theorem 4, so that  $m = n$ , and let  $M_j$  have dimension  $l = n - h$ . Let the successive minima of  $M_j$  be  $s_1 \leq \dots \leq s_l$  and choose a basis  $\mathbf{e}_1, \dots, \mathbf{e}_l$  for  $M_j$  as in Lemma 1. We now observe, according to (1.4), that there exists  $\mathbf{w} \in \mathbb{Z}^n$  and  $\rho_1, \dots, \rho_l \in \mathbb{Z}$  such that

$$\mathbf{e}_i \equiv \rho_i \mathbf{w} \pmod{Q}, \quad (1 \leq i \leq l),$$

where  $Q = q_1 \cdots q_h$ . Define  $\sigma_i = 1 + \lfloor s_i \rfloor$  for  $1 \leq i \leq l$  and consider the lattice

$$N = \{(n_1 \sigma_1, \dots, n_l \sigma_l) \in \mathbb{Z}^l : \rho_1 n_1 + \dots + \rho_l n_l \equiv 0 \pmod{Q}\}.$$

If  $\text{h.c.f.}(\rho_1, \dots, \rho_l, Q) \neq 1$  then  $M_j$  contains no primitive vectors, and hence can make no contribution to  $S(F; B, \mathbb{Z}^n)$ . We may therefore suppose that

$$\text{h.c.f.}(\rho_1, \dots, \rho_l, Q) = 1,$$

whence  $N$  has rank  $l$  and determinant

$$\det N = Q \prod_{i=1}^l \sigma_i \ll Q \det M_j,$$

by (2.2). We write  $t_1 \leq \dots \leq t_l$  for the successive minima of  $\mathbf{N}$  and choose a basis  $\mathbf{c}^{(1)}, \dots, \mathbf{c}^{(l)}$  for  $\mathbf{N}$  as in Lemma 1. Thus in particular  $|\mathbf{c}^{(j)}| = t_j$  for each  $j$  and hence

$$\prod_{j=1}^l |\mathbf{c}^{(j)}| = \prod_{j=1}^l t_j \ll \det \mathbf{N} \ll Q \det \mathbf{M}_j, \quad (7.2)$$

by a further application of (2.2). We now put

$$\mathbf{c}^{(j)} = (d_1^{(j)} \sigma_1, \dots, d_l^{(j)} \sigma_l), \quad (1 \leq j \leq l)$$

and

$$\mathbf{f}_j = Q^{-1}(d_1^{(j)} \mathbf{e}_1 + \dots + d_l^{(j)} \mathbf{e}_l), \quad (1 \leq j \leq l).$$

By construction we have  $\mathbf{f}_j \in \mathbb{Z}^n$  and  $[\mathbf{f}_j] \in M_j(\mathbb{Q})$ . Moreover, since the vectors  $\mathbf{c}^{(1)}, \dots, \mathbf{c}^{(l)} \in \mathbb{Z}^l$  are linearly independent it follows that the matrix

$$\left(d_i^{(j)}\right)_{1 \leq i, j \leq l}$$

is invertible. Bearing in mind that the vectors  $\mathbf{e}^{(1)}, \dots, \mathbf{e}^{(l)}$  are linearly independent, it follows that  $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(l)}$  are also linearly independent. Let

$$\mathbf{N}^{(*)} = \langle \mathbf{f}^{(1)}, \dots, \mathbf{f}^{(l)} \rangle$$

be the  $\mathbb{Z}$ -linear span of  $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(l)}$ , and observe that  $\det \mathbf{N}^{(*)} \ll \prod_{j=1}^l |\mathbf{f}^{(j)}|$ . Now it is clear from the triangle inequality that we have

$$|\mathbf{f}^{(j)}| \leq Q^{-1} \sum_{i=1}^l |d_i^{(j)}| |\mathbf{e}_i|,$$

for each  $1 \leq j \leq l$ . Employing (7.2) we therefore obtain

$$\begin{aligned} \det \mathbf{N}^{(*)} &\ll Q^{-l} \prod_{j=1}^l \left( \sum_{i=1}^l |d_i^{(j)}| s_i \right) \\ &\ll Q^{-l} \prod_{j=1}^l \left( \sum_{i=1}^l |d_i^{(j)}| \sigma_i \right) \\ &\ll Q^{-l} \prod_{j=1}^l |\mathbf{c}_i^{(j)}| \\ &\ll Q^{1-l} \det \mathbf{M}_j. \end{aligned}$$

Now write

$$\mathbf{M}_j^{(*)} = \{\mathbf{x} \in \mathbb{Z}^n : [\mathbf{x}] \in M_j\} \cup \{\mathbf{0}\}$$

for the lattice associated to the linear space  $M_j$ . It is worthwhile highlighting that although  $\mathbf{M}_j$  spans  $M_j$ , it is not necessarily the case that  $\mathbf{M}_j$  is equal to  $\mathbf{M}_j^{(*)}$ . However we clearly have

$$\mathbf{N}^{(*)} \subseteq \mathbf{M}_j^{(*)},$$

so that

$$H(M_j) = \det M_j^{(*)} \ll \det N^{(*)} \ll Q^{1-l} \det M_j,$$

by (7.1) and Lemma 2. We now observe that (1.5) yields

$$\det M_j \ll \prod_{i=1}^h \frac{q_i^{n-(i-1)}}{B} \ll \frac{Q^n}{B^h},$$

since we are taking  $\Lambda = \mathbb{Z}^n$ . Thus, on using the relation  $l = n - h$  we obtain

$$H(M_j) \ll Q^{1+h} B^{-h}.$$

Alternatively, since all the successive minima of  $M_j$  are  $O(B(\log \|F\| B)^{n(n+1)})$ , it follows from (2.2) that

$$\det M_j \ll B^{n-h} (\log \|F\| B)^{n(n+1)(n-h)},$$

so that

$$H(M_j) \ll Q^{1-l} B^{n-h} (\log \|F\| B)^{n(n+1)(n-h)} \ll Q^{1-n+h} B^{n-h} (\log \|F\| B)^{n^3}.$$

We use the first of these two bounds for  $Q \leq B(\log \|F\| B)^{n^2}$ , and the second otherwise. Theorem 5 then follows.

## 8 Proof of Theorem 7

Let  $S \subset \mathbb{P}^3$  denote the surface  $F = 0$ . We begin by considering the contribution to  $N(F; \mathbf{B})$  arising from a line  $L$ , defined over  $\mathbb{Q}$  and lying in the surface  $S$ . Let  $\Lambda$  be the integer lattice corresponding to  $L$ . Then  $H(L) = \det \Lambda$ , by (7.1), and we claim that

$$\begin{aligned} & \#\{\mathbf{x} \in Z^4 : [\mathbf{x}] \in L(\mathbb{Q}), |x_i| \leq B_i, (1 \leq i \leq 4)\} \\ &= \#\{\mathbf{x} \in \Lambda \cap Z^4 : |x_i| \leq B_i, (1 \leq i \leq 4)\} \\ &\ll 1 + \frac{B_1 B_2}{\det \Lambda}. \end{aligned} \tag{8.1}$$

The box we are concerned with lies inside the ellipsoid

$$\left(\frac{x_1}{B_1}\right)^2 + \left(\frac{x_2}{B_2}\right)^2 + \left(\frac{x_3}{B_2}\right)^2 + \left(\frac{x_4}{B_2}\right)^2 \leq 4.$$

Arguing in  $\mathbb{R}^4$ , the intersection of this ellipsoid with any 2-dimensional linear space will be an ellipse of area  $O(B_1 B_2)$ . Thus, to establish (8.1) it suffices to count points in  $\mathbb{R}^2$  lying in an ellipse, and which belong to a lattice  $\Lambda'$  of determinant  $\det \Lambda$ . By using an appropriate unimodular transformation  $\mathbf{M}$  we may map this ellipse to a disc of the same area. We are therefore led to count the number of lattice points contained in a disc of area  $O(B_1 B_2)$  which lie on  $\mathbf{M}\Lambda'$ . Since  $\det(\mathbf{M}\Lambda') = \det(\Lambda') = \det(\Lambda)$ , an application of [9, Lemma 1 (vi)] therefore yields the bound (8.1).

We now apply Lemma 3. If  $\log \|F\| \gg \log B_1$  this shows that we may confine attention to points on a union of  $O(1)$  curves  $C$  in  $S$ , each of degree  $O(1)$ . For those curves  $C$  of degree at least 2 we use the bound

$$N_C(B_1) \ll B_1^{1+\varepsilon},$$

which follows from Lemma 5, while for the case in which  $C$  is a line we get a satisfactory estimate from (8.1). Henceforth we may therefore assume that  $\log \|F\| \ll \log B_1$ .

By the argument used to derive Theorem 2 from Theorem 4, the points in which we are interested lie on  $O(V^{1/2}(\log B_1)^{26})$  linear subspaces of  $S$ . Those subspaces of dimension 0 are clearly satisfactory for Theorem 7, since  $V^{1/2} \leq B_1 B_2$ . Similarly those lines  $L \subset S$  which contain at most one rational point in the box under consideration also make a satisfactory contribution. The remaining lines have at least two rational points in the relevant box, and hence have height  $O(B_1^2)$ .

Recall the definition of the Fano variety  $F_1(S) = \{L \in \mathbb{G}(1, 3) : L \subset S\}$  of lines in  $S$ . Then Lemma 6 implies that the degree of  $F_1(S)$  is bounded uniformly in terms of  $d$ . Moreover, it is well-known and easy to prove that whenever  $d \geq 2$ , we have

$$\dim F_1(S) \leq 1,$$

and  $F_1(S)$  contains no linear component of dimension 1. In particular it follows from Lemma 5 that

$$N_\Phi(2H) \ll H^{1+\varepsilon},$$

for any irreducible component  $\Phi \subseteq F_1(S)$ .

Let  $L \in F_1(S)$  be a line defined over  $\mathbb{Q}$ , with height  $H < H(L) \leq 2H$  for some  $H \ll B_1^2$ . Then (8.1) shows that each such line contributes  $O(1 + B_1 B_2 / H)$  to  $N(F; \mathbf{B})$ . We have just seen that there are  $O(H^{1+\varepsilon})$  available lines. However we also know that the number of linear spaces  $M_j$  is  $O(V^{1/2} B_1^\varepsilon) = O(B_1^{1+\varepsilon} B_2)$ . It follows that the total contribution from the lines  $L = M_j \subset S$  which are defined over  $\mathbb{Q}$  and have height  $H < H(L) \leq 2H$  is

$$\ll B_1^{1+\varepsilon} B_2 + \frac{B_1 B_2}{H} \cdot H^{1+\varepsilon} = B_1 B_2 (B_1^\varepsilon + H^\varepsilon).$$

Since we have  $H(L) \ll B_1^2$  we may sum this bound over dyadic intervals with  $H \ll B_1^2$  to obtain the overall estimate  $\ll (B_1 B_2)^{1+2\varepsilon}$ . We complete the proof of Theorem 7 on re-defining  $\varepsilon$ .

## 9 Geometry of cubic threefolds

In view of the results of Heath-Brown [9, Theorem 2] and of Broberg and Salberger [2] already cited it suffices for the proof of Theorem 3 to consider the situation in which  $d = 3$ . Thus we assume that the equation  $F = 0$  defines an irreducible cubic threefold  $X \subset \mathbb{P}^4$ , throughout this section. It is clear from Theorem 4 that we must examine the possible lines and planes on cubic threefolds. For  $m = 1, 2$ , recall from §3 the Fano variety  $F_m(X) \subset \mathbb{G}(m, 4)$  of  $m$ -dimensional linear spaces contained in  $X$ . By Lemma 6 it follows that  $F_m(X)$  is an intersection of hypersurfaces whose degree and number are bounded absolutely.

We first consider the dimension of  $F_1(X)$ , since the larger this dimension is, the more difficult it will be to handle the contribution from the lines on  $X$ . Define the incidence correspondence

$$I = \{(x, L) \in X \times F_1(X) : x \in L\}.$$

The fibre of  $I$  over a point in  $F_1(X)$  has dimension 1. Hence it follows that the dimension of  $I$  is  $1 + \dim F_1(X)$ . Now consider the projection onto  $X$ . If  $x \in X$  is a generic point then the tangent hyperplane  $\mathbb{T}_x(X)$  to  $X$  at  $x$  has dimension 3. Thus the dimension of  $X \cap \mathbb{T}_x(X)$  is 2. Moreover, any line  $L \subset X$  which passes through  $x$ , must be contained in  $\mathbb{T}_x(X)$ . Hence the fibre of  $I$  over a generic point  $x \in X$  has dimension at most 1. It follows that  $\dim I \leq 4$ , and so

$$\dim F_1(X) \leq 3.$$

In a precisely similar manner a consideration of

$$J = \{(x, P) \in X \times F_2(X) : x \in P\} \quad (9.1)$$

shows that

$$\dim F_2(X) \leq 1.$$

We also record the following facts about  $F_1(X)$  and  $F_2(X)$ .

**Lemma 10.** *Let  $X \subset \mathbb{P}^4$  be an irreducible cubic threefold. Then the following hold:*

- (i) *Suppose  $\Psi \subseteq F_1(X)$  where  $\Psi$  is a plane. Then the lines  $L \in \Psi$  sweep out a plane in  $X$ .*
- (ii)  *$F_2(X)$  does not contain a line.*

*Proof.* For the proof of part (i) we observe that any plane in  $\mathbb{G}(1, 4)$  corresponds either to a locus of lines passing through the same point, which are all contained in the same hyperplane, or else to a locus of coplanar lines. In the first case we would find that  $X$  contains a hyperplane, which is impossible since  $X$  is an irreducible cubic threefold. In the second case the lines sweep out a plane, as claimed.

Similarly for part (ii) we note that a line in  $\mathbb{G}(2, 4)$  corresponds to the set of planes of  $\mathbb{P}^4$  containing a given line, and contained in a fixed hyperplane. They therefore sweep out the hyperplane. Since  $X$  does not contain a hyperplane  $F_2(X)$  cannot contain a line.  $\square$

Whenever the dimension of  $F_1(X)$  is maximal we shall call upon the following geometric result of Segre [12].

**Lemma 11.** *Let  $X \subset \mathbb{P}^4$  be an irreducible cubic threefold, and suppose that  $F_1(X)$  has dimension 3. Then*

$$\dim F_2(X) = 1.$$

Recall that any plane contains a two dimensional family of lines. Thus whenever  $\dim F_2(X) = 1$ , we see that  $X$  contains a three dimensional family of lines, all of which lie in planes contained in  $X$ . In fact most of the lines in  $X$  arise in this way, as is shown by the following result.

**Lemma 12.** *Let  $X \subset \mathbb{P}^4$  be an irreducible cubic threefold for which  $F_2(X)$  has dimension 1. Then every line in  $X$  lies in a plane contained in  $X$ , with the possible exception of those lying on a certain surface of degree  $O(1)$  contained in  $\mathbb{G}(1, 4)$ .*



Suppose firstly that  $X$  is a cone, and let  $x \in X$  be a vertex point. The family of lines in  $X$  passing through  $x$  is an algebraic variety

$$\{L \in F_1(X) : x \in L\} = Y_x,$$

say. Moreover  $Y_x$  has dimension 2 and degree  $O(1)$  by Lemma 6. Let  $L \in F_1(X) \setminus Y_x$ . Then if we write  $\overline{x, L}$  for the plane spanned by  $x$  and  $L$ , we clearly have

$$L \subset \overline{x, L} \subset X.$$

This establishes Lemma 12 whenever  $X$  is a cone.

We suppose henceforth that  $X$  is not a cone. Throughout the proof of Lemma 12, we will use  $H \in \mathbb{P}^{4*}$  to denote a generic hyperplane. Since  $X$  is not a cone, the hyperplane section

$$S_H = H \cap X$$

is an irreducible cubic surface which is not a cone. We claim that  $S_H$  contains infinitely many lines. Let  $\Phi \subseteq F_2(X)$  be an irreducible component of dimension 1, and observe that

$$X = \bigcup_{P \in \Phi} P. \quad (9.2)$$

In particular it clearly follows that

$$S_H = \bigcup_{P \in \Phi} (H \cap P).$$

The intersection  $H \cap P$  cannot ever be a plane, since  $S_H$  is an irreducible cubic surface. Hence  $H \cap P$  is always a line. Thus  $S_H$  is a union of lines and hence contains infinitely many lines as claimed.

We now call upon the following classical result, which follows from the work of Bruce and Wall [4] for example.

**Lemma 13.** *Let  $S \subset \mathbb{P}^3$  be a ruled irreducible cubic surface. Then either  $S$  is a cone or the singular locus of  $S$  is a line.*

Let  $Y \subset X$  denote the singular locus of  $X$ , and let  $T_H \subset S_H$  denote the singular locus of  $S_H$ . Since we have already seen that  $S_H$  contains infinitely many lines, Lemma 13 implies that  $T_H$  is a line. An application of Bertini's theorem (in the form given by Harris [6, Theorem 17.16], for example) therefore shows that

$$H \cap Y = T_H \cong \mathbb{P}^1$$

for  $H \in \mathbb{P}^{4*}$ . It follows that  $Y$  must be a plane. Suppose that  $X$  is given by a cubic form  $F(\mathbf{x}) = F(x_1, x_2, x_3, x_4, x_5)$ . After a linear change of variables we may assume that the plane  $Y$  is given by  $x_1 = x_2 = 0$ . It follows that there exist quadratic forms  $Q_1, Q_2$  such that

$$F(\mathbf{x}) = x_1 Q_1(\mathbf{x}) + x_2 Q_2(\mathbf{x}).$$

Upon considering the partial derivatives of  $F$  with respect to  $x_1, x_2$  one deduces further that for  $i = 1, 2$  the forms  $Q_i(0, 0, x_3, x_4, x_5)$  must vanish identically, since  $Y$  is a double plane. Hence there exist linear forms  $L_1, L_2, L_3$  such that

$$F(\mathbf{x}) = x_1^2 L_1(\mathbf{x}) + x_2^2 L_2(\mathbf{x}) + x_1 x_2 L_3(\mathbf{x}).$$

Moreover we may assume that  $L_1, L_2, L_3$  are linearly independent, since otherwise  $X$  would be a cone. We have therefore established the following result, which may be of independent interest.

**Lemma 14.** *Let  $X \subset \mathbb{P}^4$  be an irreducible cubic threefold with  $\dim F_2(X) = 1$ . Then either  $X$  is a cone or else  $X$  takes the shape*

$$x_1x_2x_3 + x_1^2x_4 + x_2^2x_5 = 0.$$

We observe that for  $X$  as in Lemma 14, there is a family of planes given by

$$\lambda x_1 - \mu x_2 = \lambda \mu x_3 + \mu^2 x_4 + \lambda^2 x_5 = 0, \quad (9.3)$$

for any  $[\lambda, \mu] \in \mathbb{P}^1$ , in addition to the plane

$$x_1 = x_2 = 0. \quad (9.4)$$

It remains to consider the lines contained in  $X$ . We hope to prove that the generic such line is contained in one of the planes (9.3). Now any line in  $\mathbb{P}^4$  can be given parametrically by

$$\mathbf{x} = [a_1s + b_1t, a_2s + b_2t, a_3s + b_3t, a_4s + b_4t, a_5s + b_5t],$$

for suitable  $a_i, b_i \in \overline{\mathbb{Q}}$ . If  $a_1b_2 - a_2b_1 = 0$  the line is contained in a hyperplane  $\lambda x_1 - \mu x_2 = 0$ , and it is readily deduced that the line lies in one of the planes (9.3) or (9.4). It therefore remains to examine the case  $a_1b_2 - a_2b_1 \neq 0$ , and here it suffices to take  $a_1 = b_2 = 1$  and  $a_2 = b_1 = 0$ . By equating coefficients in the vanishing binary form  $st(a_3s + b_3t) + s^2(a_4s + b_4t) + t^2(a_5s + b_5t)$ , we conclude that any line contained in  $X$ , which is not contained in any of the planes (9.3), must take the shape

$$\alpha x_3 = \beta x_1 + \gamma x_2, \quad \alpha x_4 = -\beta x_2, \quad \alpha x_5 = -\gamma x_1,$$

for appropriate  $[\alpha, \beta, \gamma] \in \mathbb{P}^2$  such that  $\alpha \neq 0$ . One readily finds that the family of all such lines forms a surface in  $\mathbb{G}(1, 4)$ . Any line whose Plücker coordinates do not lie on this surface will be contained in one of the planes (9.3) or (9.4). This completes the proof of Lemma 12.

It is interesting to remark that for  $X$  as in Lemma 14, the Fano variety of planes  $F_2(X)$  is the union of a single twisted cubic and an isolated point in  $\mathbb{G}(2, 4)$ . However we shall make no use of this fact in our work. We end this section by considering points which lie on infinitely many planes.

**Lemma 15.** *Let  $X$  be an irreducible cubic threefold. If  $X$  is a cone then its set of vertices is either a single point or a line. Moreover, if  $x \in X$  lies on infinitely many planes in  $X$ , then  $X$  is a cone with vertex  $x$ . Indeed if  $X$  is not a cone with vertex  $x$ , then  $x$  lies on just  $O(1)$  planes in  $X$ .*

*Proof.* If  $x$  and  $y$  are two distinct vertices of  $X$  any point  $z$  on the line  $\overline{xy}$  is also a vertex. Thus the set of vertices,  $V$  say, is necessarily a linear space. However if  $V$  has dimension 2, and  $p \in X \setminus V$ , then all of  $\overline{p, V}$  is contained in  $X$ , which is impossible, since  $X$  cannot contain a hyperplane. Now suppose that  $x \in X$  lies on infinitely many planes in  $F_2(X)$ . Let  $J$  be as defined in (9.1), so that the fibre over  $x$  is infinite, and therefore of dimension 1 since  $\dim F_2(X) \leq 1$ .

Let  $\Phi$  be an irreducible component of this fibre, with dimension 1. Then (9.2) shows that  $X$  is a cone with vertex  $x$ . Thus if  $x$  is not a vertex of  $X$  the fibre of  $J$  over  $x$ , which may be written

$$\{P \in F_2(X) : x \in P\} = Y_x,$$

say, has dimension 0. In order to complete the proof of the lemma, it therefore suffices to show that the degree of  $Y_x$  is bounded absolutely. But this follows directly from Lemma 6.  $\square$

## 10 Proof of Theorem 3

As already remarked we may restrict our attention to the case in which  $X$  has degree 3. By combining Lemma 3 and Theorem 1 we see that it suffices also to assume that  $\log \|F\| \ll \log B$ . Then, according to Theorem 4, the points in which we are interested lie on  $O(B^3(\log B)^{58})$  linear subspaces of  $X$ . Subspaces of dimension 0 are clearly satisfactory for Theorem 3, so it remains to consider the case in which  $M_j$  is a line or a plane. In view of Theorem 5 we know that  $H(M_j) \ll B(\log B)^{125}$  in these cases.

Our argument begins by considering points which lie on certain planes  $P \in F_2(X)$ . We shall call such a plane “ $B$ -good” if it contains three points  $x_1, x_2, x_3$  in general position, with  $H(x_i) \leq B(\log B)^{31}$ . We remark at once that whenever  $M_j$  is a plane it is automatically  $B$ -good if  $B$  is large enough, since Theorem 4 tells us that the successive minima of  $M_j$  are  $O(B(\log B)^{30})$ . For any  $B$ -good plane  $P$  we let  $M$  be the integer lattice corresponding to  $P$ , so that  $H(P) = \det M$ , by (7.1). Then the successive minima of  $M$  will be  $O(B(\log B)^{31})$ , so that Lemma 4 yields

$$N_P(B) \ll \prod_{j=1}^3 \left(1 + \frac{B}{s_j}\right) \ll \prod_{j=1}^3 \frac{B(\log B)^{31}}{s_j},$$

since  $B \leq B(\log B)^{31}$ . But then it follows that

$$N_P(B) \ll B^3(\log B)^{93} / \det M = B^3(\log B)^{93} / H(P).$$

We proceed to consider the contribution from all  $B$ -good planes for which  $H < H(P) \leq 2H$ . Consider those planes which belong to a particular irreducible component  $\Phi$  of  $F_2(X)$ . The number of such planes is at most  $N_\Phi(2H)$ . Moreover  $\Phi$  is either a single point, or is a curve of degree at least 2, by Lemma 10. Since the degree of  $\Phi$  is absolutely bounded, by Lemma 6, Lemma 5 shows that

$$N_\Phi(2H) \ll H^{1+\varepsilon}.$$

It follows that the total contribution from all  $B$ -good planes  $P$  with height  $H < H(P) \leq 2H$  is

$$\ll \frac{B^3(\log B)^{93}}{H} \cdot H^{1+\varepsilon}.$$

Since we have  $H(P) \leq B^3(\log B)^{93}$  we may sum this bound over dyadic intervals with  $H \ll B^3(\log B)^{93}$  to obtain an estimate  $O(B^{3+2\varepsilon})$  for the total contribution to  $N(F; B)$  arising from points on  $B$ -good planes. On re-defining  $\varepsilon$  this is

satisfactory for Theorem 3. It follows that we have a satisfactory contribution whenever  $M_j$  is a plane, and also whenever  $M_j$  is a line contained in a  $B$ -good plane.

We can now dispose of the case in which  $X$  is a cone with a line  $L$  of vertices, such that  $L$  contains two distinct rational points  $p, q$  with  $H(p), H(q) \leq B$ . In this instance, if  $x \in X \setminus L$  is a rational point with  $H(x) \ll B$ , then the plane  $\overline{p, q, x}$  lies in  $X$  and is  $B$ -good. Thus the number of such points is  $O(B^{3+\varepsilon})$  by the above. Since the line  $L$  contains  $O(B^2)$  admissible points, by Theorem 1, we may conclude that Theorem 3 holds in this case. Henceforth we shall assume that  $X$  is not of the type just considered. In the light of Lemma 15 our assumption tells us that  $X$  has at most one vertex  $v$  which is a rational point of height  $H(v) \leq B$ .

We now turn our attention to the case of lines  $M_j$ . In view of our treatment of  $B$ -good planes, it will be enough to consider lines that are not contained in any  $B$ -good plane. Let  $L$  be a line defined over  $\mathbb{Q}$ , and let  $\Lambda$  be the corresponding 2-dimensional integer lattice. Fix a basis  $\mathbf{m}_1, \mathbf{m}_2$  of  $L$  of the type given in Lemma 1. Then we shall say that  $L$  is “ $B$ -exceptional” if it does not belong to a  $B$ -good plane, and if either  $\dim F_2(X) \neq 1$ , or if  $\dim F_2(X) = 1$  and  $L$  corresponds to one of the exceptional lines in Lemma 12, or if  $\dim F_2(X) = 1$  but  $X$  is a cone with vertex  $[\mathbf{m}_1]$ . We proceed by considering the contribution from  $B$ -exceptional lines  $M_j$ . In each case we shall show that we have at most a two dimensional family of lines. If  $\dim F_2(X) \neq 1$  then  $\dim F_1(X) \leq 2$ , by Lemma 11. In the second case the claim is an immediate consequence of Lemma 12. In the third case we note that the set of lines passing through any given point of  $X$  will have dimension at most 2. The claim is therefore obvious when  $X$  has at most one vertex. If  $X$  has a line of vertices, we have seen above that there is at most one such vertex  $v$  which is defined over  $\mathbb{Q}$  and which has height  $H(v) \leq B$ . However we observe that

$$|\mathbf{m}_1| = s_1 \leq (s_1 s_2)^{1/2} \ll (\det \Lambda)^{1/2} = H(M_j)^{1/2} \ll B^{(1+\varepsilon)/2},$$

by (2.2), (7.1) and Theorem 5. Thus  $|\mathbf{m}_1| \leq B$  if  $B$  is large enough. It follows that  $[\mathbf{m}_1] = v$ , so that  $M_j$  must pass through the unique vertex  $v$  of  $X$  for which  $H(v) \leq B$ . We therefore conclude in all cases that the set of  $B$ -exceptional lines  $M_j$  has dimension at most 2.

Now it is clear that any  $B$ -exceptional line must be contained either in  $F_1(X)$  if  $\dim F_2(X) \neq 1$ , or to the closed set described in Lemma 12 if  $\dim F_2(X) = 1$  and  $L$  corresponds to one of the exceptional lines described in this result, or to the set  $\{L \in F_1(X) : [\mathbf{m}_1] \in L\}$  if  $X$  is a cone with vertex  $[\mathbf{m}_1]$ . If  $\Psi \subset F_1(X)$  is an irreducible component of any of these closed sets then we have already seen that  $\dim \Psi \leq 2$ , and it follows from Lemma 6 and Lemma 12 that  $\Psi$  has degree  $O(1)$ . Moreover, if  $\Psi$  were a plane then the  $B$ -exceptional lines  $L \in \Psi$  would all lie in a plane  $P$ , say, contained in  $X$ , by Lemma 10. The contribution from such components is therefore  $O(B^3)$ , by Theorem 1, and this is satisfactory for our theorem. Such components  $\Psi$  may therefore be ignored.

We may now complete the treatment of the  $B$ -exceptional lines in much the same way as we dealt with the  $B$ -good planes. If  $L = M_j$  is a  $B$ -exceptional line then Theorem 4 implies that the successive minima of the corresponding

lattice  $M_j$  will be  $O(B(\log B)^{30})$ , so that Lemma 4 yields

$$\begin{aligned} N_L(B) &\ll \prod_{j=1}^2 \left(1 + \frac{B}{s_j}\right) \\ &\ll \prod_{j=1}^2 \frac{B(\log B)^{30}}{s_j} \\ &\leq B^2(\log B)^{60} / \det M_j \\ &= B^2(\log B)^{60} / H(L). \end{aligned}$$

We have now to count the number of  $B$ -exceptional lines  $L$  for which  $H < H(L) \leq 2H$ . We do this by considering those which belong to a given subvariety  $\Psi \subset F_1(X)$ , where we may now restrict attention to the case in which either  $\Psi$  has dimension at most 1, or it has dimension 2 and degree at least 2. For such  $\Psi$  we see from Theorem 1 and Lemma 5 that

$$N_\Psi(2H) \ll H^{2+\varepsilon}.$$

It follows that the total contribution from all  $B$ -exceptional lines  $L$  with height  $H < H(L) \leq 2H$  is

$$\ll \frac{B^2(\log B)^{60}}{H} \cdot H^{2+\varepsilon}.$$

Since we have  $H(L) \leq B(\log B)^{125}$ , by Theorem 5, we may sum this bound over dyadic intervals with  $H \ll B(\log B)^{125}$  to obtain an estimate  $O(B^{3+2\varepsilon})$  for the total contribution to  $N(F; B)$  arising from points on  $B$ -exceptional lines. On re-defining  $\varepsilon$  this is satisfactory for Theorem 3.

It now remains to deal with the remaining lines, which we refer to as “ $B$ -ordinary”. According to our definition of  $B$ -exceptional lines, there will be no  $B$ -ordinary lines unless  $\dim F_2(X) = 1$ . Moreover, when  $\dim F_2(X) = 1$ , any  $B$ -ordinary line will be contained in a plane, by Lemma 12, but not in a  $B$ -good plane. Moreover the corresponding point  $[\mathbf{m}_1]$  will not be a vertex of  $X$ . Any  $B$ -ordinary line  $L = M_j$  with  $H < H(L) \leq 2H$  will contribute  $O(B^2/H)$  to  $N(F; B)$ , by Lemma 4, since

$$H(L) \ll B^{1+\varepsilon}$$

in Theorem 5. Thus we need to estimate the number of  $B$ -ordinary lines  $L$  with height of order  $H$ . To do this we observe that  $L$  will be spanned by the lattice generated by  $\mathbf{m}_1, \mathbf{m}_2$  with

$$H(L) \leq |\mathbf{m}_1||\mathbf{m}_2| \ll H(L),$$

by (2.2) and (7.1). In particular we will have  $|\mathbf{m}_1| \ll H^{1/2}$ , since  $s_1 \leq s_2$ . We therefore count  $B$ -ordinary lines according to the corresponding vector  $\mathbf{m}_1$ . Since  $[\mathbf{m}_1]$  is not a vertex of  $X$  it follows from Lemma 15 that there are  $O(1)$  planes through  $[\mathbf{m}_1]$ . If such a plane is not  $B$ -good, it can contain at most one line  $M_j$ . To see this we recall that a line  $M_j$  corresponds to a lattice  $M_j$  with successive minima of size  $O(B(\log B)^{30})$ . Hence, for large enough  $B$ , the line  $M_j$  will contain two distinct points of height at most  $B(\log B)^{31}$ . However, a plane that is not  $B$ -good can contain at most two linearly independent points

of such height, and hence can contain at most one line  $M_j$ , as claimed. We may therefore conclude that there are at most  $O(1)$   $B$ -ordinary lines through each point  $[\mathbf{m}_1]$ . In view of Theorem 1, the number of available  $\mathbf{m}_1$  with  $|\mathbf{m}_1| \ll H^{1/2}$  is  $O(H^2)$ , so that the total contribution from  $B$ -ordinary lines  $L$  with height  $H < H(L) \leq 2H$  is

$$\ll \frac{B^2}{H} \cdot H^2 = B^2 H \ll B^{3+\varepsilon}.$$

On summing over dyadic ranges for  $H$ , and re-defining  $\varepsilon$ , we find that this too is satisfactory for Theorem 3. This completes the proof.

## References

- [1] N. Broberg, A note on a paper by R. Heath-Brown: “The density of rational points on curves and surfaces”, *J. reine angew. Math.*, 571 (2004), 159–178.
- [2] N. Broberg and P. Salberger, Counting rational points on threefolds, 105–120. *Arithmetic of higher-dimensional algebraic varieties*, Progress in Mathematics 226, Birkhäuser, 2003.
- [3] T.D. Browning, A note on the distribution of rational points on threefolds, *Quart. J. Math.*, 54 (2003), 33–39.
- [4] J.W. Bruce and C.T. Wall, On the classification of cubic surfaces, *J. London Math. Soc.*, 19 (1979), 257–267.
- [5] H. Davenport, Cubic forms in 16 variables, *Proc. Roy. Soc. A*, 272 (1963), 285–303.
- [6] J. Harris, *Algebraic Geometry*, Springer-Verlag, 1992.
- [7] D.R. Heath-Brown, Diophantine approximation with square-free numbers, *Math. Zeit.*, 187 (1984), 335–344.
- [8] D.R. Heath-Brown, The density of rational points on cubic surfaces, *Acta Arithmetica*, 79 (1997), 17–30.
- [9] D.R. Heath-Brown, The density of rational points on curves and surfaces, *Annals of Math.*, 155 (2002), 553–595.
- [10] J. Pila, Density of integral and rational points on varieties, *Astérisque*, 228 (1995), 183–187.
- [11] W.M. Schmidt, *Diophantine Approximations and Diophantine Equations* (LNM 1467), Springer-Verlag, 1991.
- [12] B. Segre, Sulle  $V_n$  contenenti più di  $\infty^{n-k} S_k$ , I, *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat.*, 5 (1948), 193–197, II, *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat.*, 5 (1948), 275–180.