Counting Rational Points on Algebraic Varieties

D.R. Heath-Brown
Mathematical Institute, Oxford

Lecture 1—A Survey of Diophantine Equations

1.1 Introduction

In these lectures we will be interested in solutions to Diophantine equations $F(x_1, \dots, x_n) = 0$, where F is an absolutely irreducible polynomial with integer coefficients, and the solutions are to satisfy $(x_1, \dots, x_n) \in \mathbb{Z}^n$. Such an equation represents a hypersurface in $\mathbb{A}^n$, and we may prefer to talk of integer points on this hypersurface, rather than solutions to the corresponding Diophantine equation. In many cases of interest the polynomial F is homogeneous, in which case the equation defines a hypersurface in $\mathbb{P}^{n-1}$, and the non-zero integer solutions correspond to rational points on this hypersurface. In this situation the solutions of $F(x_1, \dots, x_n) = 0$ form families of scalar multiples, and each family produces a single rational point on the corresponding projective hypersurface. Occasionally we shall encounter systems of two or more equations, and these may correspond to varieties of codimension 2 or more, rather than hypersurfaces.

Much work in the theory of Diophantine equations has been directed at showing that certain classes of equations have finitely many solutions. However we shall be interested in those cases where either we expect the number of solutions to be infinite, or we expect the number to be finite but cannot prove it. In these cases it is sensible to ask for bounds on the number of solutions which might lie in a large region $\max |x_i| \leq B$, say.

1.2 Examples

Let us look at some examples.

1. The equation
$$x_1^k + x_2^k = x_3^k + x_4^k. \tag{1.1}$$
It is expected that there are only the trivial solutions as soon as $k \geq 5$, but since we are unable to prove this, one may ask for an upper bound on the number of non-trivial solutions in a given large box.

2. The equation
$$x_1^k + x_2^k + x_3^k = N, \quad x_1, x_2, x_3 \geq 0.$$
Here it is believed that there is at most one solution, up to permutation, as soon as $k \geq 7$, but in the absence of a proof we ask for upper bounds on the number of solutions. (When $k \leq 6$ we know of infinitely many essentially different examples in which N has two or more representations.)

3. In Waring's problem one encounters the equation

$$x_1^k + \ldots + x_s^k = x_{s+1}^k + \ldots + x_{2s}^k, \quad 0 \le x_1, \ldots, x_{2s} \le B.$$

If one can show that there are $O(B^{2s-k+\varepsilon})$ solutions, for any fixed $\varepsilon > 0$, one can deduce the Hardy-Littlewood asymptotic formula for representations of a large integer N as a sum of $2s+1$ perfect k-th powers. Thus one would have $G(k) \le 2s+1$, providing s is large enough for the usual local conditions to hold.

4. Vinogradov's Mean Value Theorem relates to the system of equations

$$x_1^h + \ldots + x_s^h = x_{s+1}^h + \ldots + x_{2s}^h, \quad (1 \le h \le k),$$

in which $0 \le x_1, \ldots, x_{2s} \le B$. It is known that if s is sufficiently large, then the number of solutions is $O(B^{2s-k(k+1)/2})$. (In fact this is known to hold if $s \ge \{1 + o(1)\}k^2 \log k$.) Such bounds have numerous applications, for example to estimates for the zero-free region of the Riemann Zeta-function. One could conjecture that the same bound holds as soon as $s > k(k+1)/2$. If true, this would lead to improved results on the Zeta-function.

5. Manin's Conjecture. As a simple special case of Manin's conjecture, let $F(x_1, x_2, x_3, x_4)$ be a non-singular[1] cubic form with integral coefficients, and suppose that there is at least one non-zero integral solution to the equation

$$F(x_1, x_2, x_3, x_4) = 0.$$

Then the conjecture states that the number of non-trivial solutions in the box $\max |x_i| \le B$ will be asymptotically $cB(\log B)^r$ for a suitable positive constant c, where r is the rank of the Picard group of the surface $F = 0$. (This is not quite the usual formulation, since we have not insisted that our solutions should be projectively distinct.) No non-singular cubic surface is known for which Manin's conjecture can be established. Indeed, even the weaker statement that the number of non-trivial solutions is $O(B^{1+\varepsilon})$ for every $\varepsilon > 0$, eludes us.

6. For $D > 0$ the equation

$$x^2 + Dy^2 = z^3, \quad 1 \le z \le D^{1/2}$$

may be used to count ideal classes of order 3 in the class group of $\mathbb{Q}(\sqrt{-D})$. It is conjectured that the number of solutions, and hence the number of such ideal classes, should be $O(D^{\varepsilon})$ for any $\varepsilon > 0$, but to date we cannot reduce the exponent below $1/2$. This question is related to the problem of giving an upper bound for the number of rational elliptic curves with given conductor.

7. It is conjectured that any irreducible polynomial $f(X) \in \mathbb{Z}[X]$ which satisfies the obvious congruence conditions should assume infinitely many square-free values. This has been established only for polynomials f of

[1] A form in n variables will be said to be non-singular if $\nabla F(\mathbf{x})$ is non-zero for every non-zero $\mathbf{x} \in \overline{\mathbb{Q}}^n$.

degree at most 3. What is required for further progress is a good bound for the number of solutions of the equation

$$f(x) = y^2 z, \quad 1 \le x \le N, \ \ y \ge N.$$

These examples demonstrate that the general problem under consideration underlies a very diverse range of questions in number theory. Although many of the above examples involve inhomogeneous equations, we shall begin by considering only the case in which F is a form. Later on we shall see how the inhomogeneous case can be handled in an analogous way to the homogeneous one. We therefore state formally the following assumptions for future reference.

Convention *Unless stated explicitly otherwise, we shall write* $\mathbf{x}$ *for the vector* $(x_1, \dots, x_n)$ *and assume that* $F(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$ *is an absolutely irreducible form of degree* d.

1.3 The Heuristic Bounds

It will be convenient to define

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

Recall that F has total degree d. Then for the vectors $\mathbf{x}$ under consideration, the values $F(\mathbf{x})$ will all be of order B^d, and indeed a positive proportion of them will have exact order B^d. Thus the 'probability' that a randomly chosen value of $F(\mathbf{x})$ should vanish might be expected to be of order B^{-d}. Since the number of vectors $\mathbf{x}$ to be considered has order B^n, this heuristic argument leads one to expect that $N^{(0)}(B)$ is of exact order B^{n-d}.

Clearly there are many things wrong with this argument, not least the fact that when $n < d$ even one solution to $F(\mathbf{x}) = 0$ would show that $N^{(0)}(B) \not\to 0$. However we can safely summarize things the following way.

Heuristic Expectation *When* $n \ge d$ *we have*

$$B^{n-d} \ll N^{(0)}(B) \ll B^{n-d}$$

unless there is a reason why not!

Certainly local (congruence) conditions will often provide a reason why $N^{(0)}(B)$ might be identically zero. However, in support of the above heuristic argument we have the following very general result of Birch.

Theorem 1 *Suppose* $F(\mathbf{x})$ *is a non-singular form of degree* d *in* $n > 2^d(d-1)$ *variables. Then there is a constant* $c_F > 0$ *such that*

$$N^{(0)}(B) \sim c_F B^{n-d},$$

providing that $F(\mathbf{x}) = 0$ *has non-trivial solutions in* $\mathbb{R}$ *and each p-adic field* $\mathbb{Q}_p$.

Since forms are in general non-singular, Birch's result answers our question completely for typical forms with $n > 2^d(d-1)$. It would be of considerable interest to reduce the lower bound for n, but except for $d \le 3$ this has not been done. In view of Birch's theorem our interest will be centred on the case in which n is small compared with d.

As has been mentioned there are many cases in which $N^{(0)}(B)$ is not of order B^{n-d}. A good illustration is provided by the diagonal cubic equation

$$x_1^3 + x_2^3 + x_3^3 + x_4^3 = 0.$$

Here there are 'trivial' solutions of the type $(a, -a, b, -b)$ which already contribute $\gg B^2$ to $N^{(0)}(B)$.

If we use the counting function $N^{(0)}(B)$ we see that a single non-zero solution $F(\mathbf{x}_0) = 0$ will produce $\ll B$ scalar multiples, so that $N^{(0)}(B) \gg B$. This behaviour often masks the contribution of other solutions. Thus it is usually convenient to count only primitive solutions, where a non-zero vector $(x_1, \dots, x_n)$ is said to be primitive if h.c.f.$(x_1, \dots, x_n) = 1$. Indeed since the vector $-\mathbf{x}_0$ will be a solution of $F(\mathbf{x}) = 0$ if and only if $\mathbf{x}_0$ is also a solution, it is natural to define

$$N(B) = N(F; B) = \frac{1}{2}\#\{\mathbf{x} \in \mathbb{Z}^n : \ F(\mathbf{x}) = 0, \ \text{h.c.f.}(x_1, \dots, x_n) = 1, \ \max|x_i| \le B\}. \quad (1.2)$$

1.4 Curves

When F is homogeneous and $n = 3$ the equation $F(x_1, x_2, x_3) = 0$ describes a projective curve in $\mathbb{P}^2$. In this situation a great deal is known. Such a curve has a genus g which is an integer in the range $0 \le g \le (d-1)(d-2)/2$. The generic curve of degree d will have $g = (d-1)(d-2)/2$.

When $g = 0$ the curve either has no rational points (as for example, when $F(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$) or it can be parameterized by rational functions. Such a parameterization allows us to estimate $N(B)$. For example, when $F(\mathbf{x}) = x_1^2 + x_2^2 - x_3^2$ the solutions take the form

$$(x_1, x_2, x_3) = \big(a(b^2 - c^2), 2abc, a(b^2 + c^2)\big) \quad \text{or} \quad \big(2abc, a(b^2 - c^2), a(b^2 + c^2)\big).$$

It is then easy to see that $N(B)$ is precisely of order B^{n-d}, since $n - d = 1$ in this case.

On the other hand, a second example with genus $g = 0$ is provided by the cubic curve $x_1^2 x_2 - x_3^3 = 0$. In this case the solutions are proportional to $(a^3, b^3, a^2 b)$. One therefore sees that $N^{(0)}(B)$ is precisely of order $B^{2/3}$. In this example we have $n - d = 0 < 2/3$.

We now turn to curves of genus 1. Either such a curve has no rational points or it is an elliptic curve. In the latter case the set of rational points can be given an abelian group structure, and the Mordell-Weil Theorem tells us that the group has finite rank r, say. Moreover, it can be shown using Néron's theory of heights, that

$$N(B) \sim c_F (\log B)^{r/2} \quad (1.3)$$

where c_F is a non-zero constant. A cubic curve with a rational point is an elliptic curve, and in this case $n - d = 0$ so that $N(B)$ grows faster than B^{n-d} as soon as $r \ge 1$.

Finally we consider curves of genus $g \ge 2$. Here the celebrated theorem of Faltings [8] shows that there are finitely many rational points, so that

$$N(B) \ll_F 1. \quad (1.4)$$

1.5 Surfaces

We have already seen the example

$$x_1^3 + x_2^3 + x_3^3 + x_4^3 = 0 \tag{1.5}$$

in which the 'trivial' solutions already contribute $\gg B^2$ to $N^{(0)}(B)$. These trivial solutions satisfy the conditions $x_1 + x_2 = x_3 + x_4 = 0$, or $x_1 + x_3 = x_2 + x_4 = 0$, or $x_1 + x_4 = x_2 + x_3 = 0$. In each case the trivial solutions are those that lie on certain lines in $\mathbb{P}^3$. These lines lie in the surface (1.5), since the equations $x_1 + x_2 = x_3 + x_4 = 0$, for example, imply $x_1^3 + x_2^3 + x_3^3 + x_4^3 = 0$.

In general, we shall *define* a trivial solution to $F(x_1, x_2, x_3, x_4) = 0$ to be one that lies on a line which is contained in the corresponding surface. Moreover we shall define $N_1(B)$ to be the counting function analogous to (1.2), but in which only non-trivial solutions are counted. Thus in the case $d = 3$, Manin's conjecture predicts the behaviour of $N_1(B)$.

In the example (1.5) we know a complete parametric solution, due to Euler, giving all the rational points as

$$\begin{aligned} x_1 &= (3b - a)(a^2 + 3b^2)c + c^4, \\ x_2 &= (3b + a)(a^2 + 3b^2)c - c^4, \\ x_3 &= (a^2 + 3b^2)^2 - (3b + a)c^3, \\ x_4 &= -(a^2 + 3b^2)^2 - (3b - a)c^3. \end{aligned} \tag{1.6}$$

Although this produces all the rational points, it is unfortunately the case that the values of $x_1, \dots, x_4$ may be integers with a large common factor, even when h.c.f.$(a, b, c) = 1$. In the absence of any good way to control such a common factor, Euler's formula is rather little use in producing an upper bound for $N_1(B)$. Indeed if one wishes to produce a lower bound, the obvious procedure is to use integral values $a, b, c \ll B^{1/4}$. In this way one can at best show that $N_1(B) \gg B^{3/4}$, while other methods yield lower bounds $N_1(B) \gg B$ and better. Thus even a complete parameterization of the solutions does not solve our problem.

A second instructive example is provided by the equation

$$x_1^4 + x_2^4 = x_3^4 + x_4^4 \tag{1.7}$$

Here there is a family of non-trivial solutions given (also by Euler) as

$$\begin{aligned} x_1 &= a^7 + a^5b^2 - 2a^3b^4 + 3a^2b^5 + ab^6, \\ x_2 &= a^6b - 3a^5b^2 - 2a^3b^4 + a^2b^5 + b^7, \\ x_3 &= a^7 + a^5b^2 - 2a^3b^4 - 3a^2b^5 + ab^6, \\ x_4 &= a^6b + 3a^5b^2 - 2a^3b^4 + a^2b^5 + b^7. \end{aligned} \tag{1.8}$$

Not all solutions have this form, but these suffice on taking $a, b \ll B^{1/7}$ to show that $N_1(B) \gg B^{2/7}$. (There is the primitivity condition on $\mathbf{x}$ to be dealt with, but this can be satisfactorily handled.) As a, b run over all possible values, the corresponding vectors $\mathbf{x}$ run over a curve in the surface (1.7). Indeed, since the curve is parameterized, it is a curve of genus zero. In this example therefore we see that a surface may contain a large number of points by virtue of there being a genus zero curve lying in the surface.

A related example is the Euler surface

$$x_1^4 + x_2^4 + x_3^4 = x_4^4.$$

Here it was shown by Elkies [4] that there is a genus 1 curve of positive rank lying in the surface. In view of Néron's result (1.3) this shows that $N_1(B) \gg (\log B)^{1/2}$. Thus we see that surfaces may contain infinitely many points when there is a curve Γ of genus one on the surface, such that Γ itself has infinitely many points.

In general we expect that this sort of behaviour is essentially all that can occur. The following is a consequence of a conjecture of Lang.

Conjecture 1 *A surface of general type contains only finitely many curves of genus zero or one, and contains only finitely many rational points not on one of these curves.*

The definition of "general type" is somewhat technical. However we note that a non-singular surface in $\mathbb{P}^3$ will be of general type as soon as $d \geq 5$.

1.6 Higher Dimensions

For varieties of higher dimension there are analogous phenomena, and we have the following conjecture, which is again a consequence of Lang's conjecture.

Conjecture 2 *On a variety of general type all rational points belong to one of a finite number of proper subvarieties.*

Lecture 2—A Survey of Results

In the remainder of these notes we shall allow all the constants implied by the $\ll$ and $O(\ldots)$ notations to depend on the degree d of the form F and on the number n of variables. However, where there is any further dependence on F we shall say so explicitly.

In this lecture, where results are formally stated as theorems, they will either be proved in full in the lectures that follow, or may be found in the author's paper [15].

2.1 Early Approaches

Until recently there have been few general results giving bounds for $N(F;B)$. Perhaps the first, historically, is due to Cohen, in the appendix to the lecturer's paper [11], where it is shown that

$$N(F;B) \ll_{\varepsilon,F} B^{n-3/2+\varepsilon}$$

for any $\varepsilon > 0$, as soon as $d \geq 2$. The proof uses the large sieve, and information of the behaviour of F modulo many different primes.

A second approach uses exponential sums to a fixed modulus, the latter being chosen to have size a suitable power of B. To work effectively the method requires F to be non-singular. One can then show (Heath-Brown [12]) that

$$N(F;B) \ll_{\varepsilon,F} B^{n-2+2/(n+1)+\varepsilon}$$

for $d \geq 2$, and

$$N(F;B) \ll_{\varepsilon,F} B^{n-3+15/(n+5)+\varepsilon} \tag{2.1}$$

for $d \geq 3$, again for any $\varepsilon > 0$. This latter result yields the estimate

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

for non-singular F of degree $d \geq 2$, as soon as $n \geq 10$.

Hooley [17] uses a sieve method rather different from Cohen's, which can be coupled with estimates for multi-dimensional exponential sums. Although no general results have been worked out, the method is quite efficient in those special cases for which it has been used. Thus for the equations (1.1), Hooley shows [18], [19] and [21] that one has

$$N_1(B) \ll_{\varepsilon,k} B^{5/3+\varepsilon}$$

when $k \geq 3$.

Other general methods depend on elementary differential geometry, as in Schmidt [29]. These techniques improve slightly on Cohen's result, and apply also to certain non-algebraic hypersurfaces.

2.2 The Method of Bombieri and Pila

The most successful general method appears to be that introduced by Bombieri and Pila [2], and developed by the lecturer [15]. In their original work, Bombieri and Pila showed that if $f(x,y) \in \mathbb{Z}[x,y]$ is an absolutely irreducible polynomial of degree d, then

$$\#\{(x,y) \in \mathbb{Z}^2 : f(x,y) = 0, |x|, |y| \leq B\} \ll_{\varepsilon} B^{1/d+\varepsilon}. \tag{2.2}$$

Indeed their result was slightly more precise than this. One very important feature of this result is that it is completely uniform with respect to f.

The estimate (2.2) is essentially best possible, as the example $f(x,y) = x^d - y$ shows. Here there are $\ll B^{1/d}$ solutions $x = m$, $y = m^d$ with $m \ll B^{1/d}$.

Pila went on [28] to apply (2.2) to our general setting, and showed that

$$N(B) \ll_{\varepsilon} B^{n-2+1/d+\varepsilon}. \tag{2.3}$$

In the case of quadratic forms one can do better, and an elementary argument shows that

$$N(B) \ll_{\varepsilon} B^{n-2+\varepsilon}$$

if $d = 2$, see [15, Theorem 2]. However it is an interesting open question whether one can extend this to higher degree forms.

Conjecture 3 *For given $d \geq 3$ and $n \geq 3$ we have*

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

It would even be interesting to know whether this could be achieved with a possible dependence on F in the implied constant. The bound (2.1) achieves this for non-singular forms F, when $n \geq 10$.

The arguments of Bombieri and Pila [2], and of Pila [28] were essentially affine in nature, and only used properties of $\mathbb{A}^2$. This left open the question of whether there might be a natural extension to higher dimensions. In particular Pila's work [28] only used the bound (2.2), applying it to plane sections of the variety $F = 0$.

2.3 Projective Curves

The lecturer's work [15] extended the method of Bombieri and Pila to projective hypersurfaces of arbitrary dimension. We shall begin by describing the result obtained for curves.

Theorem 2 *Let $F(x_1, x_2, x_3) \in \mathbb{Z}[x_1, x_2, x_3]$ be an absolutely irreducible form of degree d, and let $\varepsilon > 0$. Then*

$$N(F; B) \ll_\varepsilon B^{2/d+\varepsilon}. \tag{2.4}$$

At first sight this is uninteresting, since it is clearly surpassed by the results of Néron (1.3) and Faltings (1.4). The difference however lies in the fact that (2.4) is uniform in F. If one tries to adapt the proof of (1.3), say, to investigate uniformity in F one finds that the rank of the curve comes into play. At present we have insufficient information about the size of the rank of elliptic curves to produce unconditional bounds, so the approach fails. None the less the lecturer has shown [14] that one has

$$N(F; B) \ll_\varepsilon B^\varepsilon$$

for non-singular cubic curves, under the assumption of the Birch-Swinnerton-Dyer conjecture and the Riemann Hypothesis for L-functions of elliptic curves. The second comment that must be made in relation to (2.4) is that it applies to curves of genus zero as well as to curves of genus one or more. If one looks at the example $F(\mathbf{x}) = x_1^d - x_2^{d-1}x_3$, then one sees that there are solutions $(m^{d-1}n, m^d, n^d)$, and these suffice to show that $N(B) \gg B^{2/d}$. It follows that (2.4) is essentially best possible. Moreover the exponent $2/d$ is clearly a considerable improvement on the value $1 + 1/d$ which the bound (2.3) would produce. None the less, the fact remains that the proof of Theorem 2 works only with the degree d, and fails to distinguish the genus of the curve. This is a serious defect in the approach.

In fact we need not require F to be absolutely irreducible in Theorem 2. Indeed for forms which are irreducible over $\mathbb{Q}$ but reducible over $\overline{\mathbb{Q}}$ we have the following stronger estimate.

Theorem 3 *Let $F(x_1, x_2, x_3) \in \overline{\mathbb{Q}}[x_1, x_2, x_3]$ be an absolutely irreducible form of degree d, but not a multiple of a rational form. Then $N(F; B) \leq d^2$. Moreover, if $F(x_1, x_2, x_3) \in \mathbb{Z}[x_1, x_2, x_3]$ is a form of degree d which is irreducible over $\mathbb{Q}$ but not absolutely irreducible, then $N(F; B) \leq d^2$.*

To prove the first assertion one writes F as a linear combination $\sum \lambda_i F_i$ of rational forms F_i, with linearly independent λ_i. Some F_i is not a multiple of F, but all rational zeros of F must satisfy $F = F_i = 0$. The result then follows by Bézout's Theorem. The second statement clearly follows from the first, on splitting F into its irreducible factors over $\overline{\mathbb{Q}}$.

We can also estimate the number of points on curves in $\mathbb{P}^3$.

Theorem 4 *Let C be an irreducible curve in $\mathbb{P}^3$, of degree d, not necessarily defined over the rationals. Then C has $O_\varepsilon(B^{2/d+\varepsilon})$ primitive points $\mathbf{x} \in \mathbb{Z}^4$ in the cube $\max |x_i| \leq B$.*

This can be established by projecting C onto a suitable plane, and counting the points on the resulting plane curve. In general such a projection may have degree less than d. However, the generic projection has degree equal to d, so that the result follows providing that one chooses the projection map with suitable care.

As with the result of Bombieri and Pila, all the above bounds are completely independent of F. The key to this is the following result, in which we write $||F||$ for the height of the form F, defined as the maximum modulus of the coefficients of F.

Theorem 5 *Suppose that $F(x_1, x_2, x_3) \in \mathbb{Z}[\mathbf{x}]$ is a non-zero form of degree d, and that the coefficients of F have no common factor. Then either $N(F; B) \leq d^2$ or $||F|| \ll B^{d(d+1)(d+2)/2}$.*

Thus, if one has a bound of the shape

$$N(F; B) \ll_\varepsilon B^{\theta+\varepsilon}||F||^\varepsilon,$$

valid for any $\varepsilon > 0$, then one can deduce that either $N(F; B) \ll 1$ or

$$N(F; B) \ll_\varepsilon B^{\theta+\varepsilon} B^{d(d+1)(d+2)\varepsilon/2}.$$

On re-defining ε we see in either case that

$$N(F; B) \ll_\varepsilon B^{\theta+\varepsilon}.$$

Thus the dependence on $||F||$ miraculously disappears!

The results described here should be compared with those in the work of Elkies [5]. The emphasis in [5] is on algorithms for searching for rational points. Elkies shows in [5, Theorem 3] that one can find the rational points of height at most B on a curve C of degree d, in time $O_{C,\varepsilon}(B^{2/d+\varepsilon})$. Thus in particular there are $O_{C,\varepsilon}(B^{2/d+\varepsilon})$ points to be found. Uniformity in C is not considered, but it seems quite plausible that the methods may yield a good dependence on the height of C, or even complete independence as in the theorems quoted above. The techniques used in the two papers show interesting similarities, although the precise relationship remains unclear. In fact Elkies also looks at surfaces and varieties of higher dimension, and gives a heuristic argument that leads to the same exponent $3/\sqrt{d}$ as occurs in Theorems 7 and 11 below.

2.4 Surfaces

In the previous lecture it was explained that the natural counting function for surfaces should exclude trivial solutions, and the function $N_1(B)$ was introduced. If there is a line L, defined over $\mathbb{Q}$, and lying in the surface $F(\mathbf{x}) = 0$, then points on L will contribute $\gg_F B^2$ to $N(F; B)$.

In analogy with Conjecture 3 we may now expect the following.

Conjecture 4 *Let $F(x_1, x_2, x_3, x_4) \in \mathbb{Z}[x_1, x_2, x_3, x_4]$ be an absolutely irreducible form of degree d, and let $\varepsilon > 0$. Then*

$$N_1(F; B) \ll_\varepsilon B^{1+\varepsilon}.$$

This would be best possible, as the example

$$x_1^d + x_2^d - x_2^{d-2}x_3x_4 = 0$$

shows. This surface is absolutely irreducible, and contains no lines other than those in the planes $x_2 = 0$, $x_3 = 0$ and $x_4 = 0$. However there are rational points $(0, ab, a^2, b^2)$, which yield $N_1(B) \gg B$.

That any points on lines in the surface will dominate the function $N(B)$ is shown by the following result.

Theorem 6 *For an absolutely irreducible form $F(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_4]$ of degree 3 or more, we have*

$$N_1(F; B) \ll_\varepsilon B^{52/27+\varepsilon}.$$

Moreover we can improve substantially on this for large values of d, as follows.

Theorem 7 *Let $F(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_4]$ be an absolutely irreducible form of degree d. Then we have*

$$N_1(F; B) \ll_\varepsilon B^{1+3/\sqrt{d}+\varepsilon}.$$

Surfaces of the type $G(x_1, x_2) = G(x_3, x_4)$, where G is a binary form, have been investigated fairly extensively in the past, although success has been limited to the cases in which $d = 3$ (Hooley [16], [22]), or $d = 4$ and G has the shape $ax^4+bx^2y^2+cy^4$ (Hooley [20]), or G is diagonal (Bennet, Dummigan and Wooley [1]). In the first two cases the methods save only a power of $\log B$ relative to B^2.

The above mentioned works were designed to show that almost all integers represented by G have essentially only one representation. We can now prove this for arbitrary irreducible forms. To formulate this precisely, we define an automorphism of the binary form G to be an invertible 2×2 matrix M, such that $G(M\mathbf{x}) = G(\mathbf{x})$ identically in $\mathbf{x}$. We then say that integral solutions of $G(\mathbf{x}) = n$ are equivalent if and only if they are related by such an automorphism with a rational matrix M. (One slightly strange consequence of this definition is that when $d = 1$ or $d = 2$ all non-zero integer solutions of $G(\mathbf{x}) = n$ are equivalent.)

We then have the following result.

Theorem 8 *Let $G(x, y) \in \mathbb{Z}[x, y]$ be a binary form of degree $d \geq 3$, irreducible over $\mathbb{Q}$. Then G has $O_d(1)$ automorphisms. Moreover the number of positive integers $n \leq X$ represented by the form G is of exact order $X^{2/d}$, providing that G assumes at least one positive value. Of these integers n there are $O_{\varepsilon,G}(X^{52/(27d-2)+\varepsilon})$ for which there are two or more inequivalent integral representations.*

The statement that the number of representable integers is of exact order $X^{2/d}$ is a classical result of Erdős and Mahler [7], dating from 1938.

Although Theorem 6 shows that points on lines may predominate, it does not automatically verify Conjecture 3 for surfaces, since there may be infinitely many lines. However this possibility can he handled successfully, and we have the following result.

Theorem 9 *Let $F(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_4]$ be an absolutely irreducible form of degree $d \geq 2$. Then we have*

$$N(F; B) \ll_\varepsilon B^{2+\varepsilon}.$$

When F is non-singular we can do better than Theorem 6, and we have the following results.

Theorem 10 *Let $F(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_4]$ be a non-singular form of degree $d \geq 2$. Then we have*

$$N_1(F; B) \ll_\varepsilon B^{4/3+16/9d+\varepsilon}. \tag{2.5}$$

Theorem 11 *Let $F(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_4]$ be a non-singular form of degree $d \geq 2$. Then we have*

$$N_1(F; B) \ll_\varepsilon B^{1+\varepsilon} + B^{3/\sqrt{d}+2/(d-1)+\varepsilon}. \tag{2.6}$$

In particular

$$N_1(F; B) \ll_\varepsilon B^{1+\varepsilon}, \tag{2.7}$$

when $d \geq 13$. Let $N_2(F; B)$ be the number of points counted by $N(F; B)$, but not lying on any curve of degree $\leq d-2$ contained in the surface. Then

$$N_2(F; B) \ll_\varepsilon B^{3/\sqrt{d}+2/(d-1)+\varepsilon}. \tag{2.8}$$

Let $N_3(F; B)$ be the number of points counted by $N(F; B)$, but not lying on any genus zero curve of degree $\leq d-2$ contained in the surface. Then

$$N_3(F; B) \ll_{\varepsilon,F} B^{3/\sqrt{d}+2/(d-1)+\varepsilon}. \tag{2.9}$$

In particular we see that Conjecture 4 holds for $d \geq 13$, providing that F is non-singular. We note that the exponent in Theorem 10 is better for $d = 3, 4$ and 5, but that otherwise one should use Theorem 11.

It is plain that curves of low degree lying in the surface $F = 0$ are a potential source for a large contribution to $N(F; B)$. Thus the following geometric result, due to Colliot-Thélène, is of great significance.

Theorem 12 *[15, Theorem 12] Let $F(\mathbf{x}) = 0$ be a non-singular surface in $\mathbb{P}^3$, of degree d. Then for every degree $\delta \leq d-2$ there is a constant $N(\delta, d)$, independent of F, such that the surface $F(\mathbf{x}) = 0$ contains at most $N(\delta, d)$ irreducible curves of degree δ.*

When $d = 3$ we have the familiar fact that a non-singular cubic surface has 27 lines. We can therefore take $N(1, 3) = 27$.

We now see from the estimate (2.9) that, with very few exceptions, the rational points on a non-singular surface of large degree are restricted to a finite number of curves of genus zero. This may be compared with the assertion of Conjecture 1.

The special diagonal surfaces

$$F(\mathbf{x}) = x_1^d + x_2^d - x_3^d - x_4^d = 0 \tag{2.10}$$

have received a great deal of attention, and it has been shown that

$$N_1(B) \ll B^{4/3+\varepsilon} \quad (d = 3)$$

(Heath-Brown [13]),

$$N_1(B) \ll B^{5/3+\varepsilon} \quad (4 \leq d \leq 7) \tag{2.11}$$

(Hooley [19] and [21]), and

$$N_1(B) \ll B^{3/2+1/(d-1)+\varepsilon} \quad (d \geq 8)$$

(Skinner and Wooley [30]). Theorem 11 supersedes these as soon as $d \geq 6$. However Browning [3] has recently shown that (2.6) may be replaced by

$$N_1(F; B) \ll_\varepsilon B^{2/3+\varepsilon} + B^{3/\sqrt{d}+2/(d-1)+\varepsilon}$$

for these particular surfaces. We shall improve this further as follows.

Theorem 13 *When $d \geq 8$ the surface (2.10) contains no genus zero curves other than the lines. Hence*

$$N_1(F; B) \ll_\varepsilon B^{3/\sqrt{d}+2/(d-1)+\varepsilon}$$

for any $d \geq 2$.

2.5 A General Result

We may now state the key result underlying most of the estimates described in this lecture. It applies to projective hypersurfaces of arbitrary dimension.

Theorem 14 *Let $F(x_1, \dots, x_n) \in \mathbb{Z}[\mathbf{x}]$ be an absolutely irreducible form of degree d, and let $\varepsilon > 0$ and $B \geq 1$ be given. Then we can find $D = D(n, d, \varepsilon)$ and an integer k with*

$$k \ll_{n,d,\varepsilon} B^{(n-1)d^{-1/(n-2)}+\varepsilon} (\log ||F||)^{2n-3},$$

as follows. There are forms $F_1(\mathbf{x}), \dots, F_k(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$, coprime to $F(\mathbf{x})$ and with degrees at most D, such that every point $\mathbf{x}$ counted by $N(F; B)$ is a zero of some form $F_j(\mathbf{x})$.

One should note that it is crucial for the degrees of the forms F_j to be suitably bounded, since one can easily construct a form $F_1(\mathbf{x})$, with degree dependent on B, which vanishes at every integer vector in the cube $\max |x_i| \leq B$.

In the case $n = 3$, Theorem 14 shows that every point counted by $N(F; B)$ satisfies $F(\mathbf{x}) = F_j(\mathbf{x}) = 0$ for some $j \ll_\varepsilon B^{2/d+\varepsilon} (\log ||F||)^3$. Bézout's Theorem shows that there are at most dD points for each j, so that $N(F; B) \ll_\varepsilon B^{2/d+\varepsilon} (\log ||F||)^3$. Theorem 2 then follows via an application of Theorem 5.

For the case $n = 4$ we see in the same way that the relevant points will lie on $O_\varepsilon(B^{3/\sqrt{d}+\varepsilon} (\log ||F||)^5)$ curves in the surface $F = 0$, each curve having degree at most dD. We may apply Theorem 4 to estimate the number of points on such a curve, but it is useful to have further information on the possible degrees of such curves. This is provided by Theorem 12 when the form F is non-singular, but otherwise we merely use the fact that the degree of the curve will be at least two for points counted by $N_1(F; B)$. In this way we may establish Theorems 7 and 11, after treating the factor $(\log ||F||)^5$ through a version of the process employed for Theorem 5.

2.6 Affine Problems

Although our main emphasis has been on integer zeros of forms, one can successfully tackle problems involving general (inhomogeneous) polynomials. For this section we therefore suppose that $F(x_1, \dots, x_n) \in \mathbb{Z}[\mathbf{x}]$ is an absolutely irreducible polynomial of total degree d, and we consider

$$N(F; B_1, \dots, B_n) = N(F; \mathbf{B}) = \\ \#\{\mathbf{x} \in \mathbb{Z} : F(\mathbf{x}) = 0,\ |x_i| \le B_i,\ (1 \le i \le n)\} \tag{2.12}$$

where $B_i \ge 1$ for $1 \le i \le n$. Thus the Bombieri-Pila result (2.2) shows that $N(F; B, B) \ll_\varepsilon B^{1/d+\varepsilon}$. In applications it can be very useful to allow the B_i to have varying sizes. Indeed one can formulate the homogeneous problem with general boxes rather than cubes, and prove an extension of Theorem 14. In our case we shall see that the following analogue of Theorem 14 holds.

Theorem 15 *Let $F(x_1, \dots, x_n) \in \mathbb{Z}[\mathbf{x}]$ be an absolutely irreducible polynomial of degree d, and let $\varepsilon > 0$ and $B_1, \dots, B_n \ge 1$ be given. Define*

$$T = \max\{\prod_{i=1}^{n} B_i^{e_i}\},$$

where the maximum is taken over all integer n-tuples $(e_1, \dots, e_n)$ for which the corresponding monomial

$$x_1^{e_1} \dots x_n^{e_n}$$

occurs in $F(\mathbf{x})$ with non-zero coefficient.

Then we can find $D = D(n, d, \varepsilon)$ and an integer k with

$$k \ll_{n,d,\varepsilon} T^\varepsilon \exp\{(n-1)(\frac{\prod \log B_i}{\log T})^{1/(n-1)}\}(\log ||F||)^{2n-3},$$

as follows. There are polynomials $F_1(\mathbf{x}), \dots, F_k(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$, coprime to $F(\mathbf{x})$ and with degrees at most D, such that every point $\mathbf{x}$ counted by $N(F; \mathbf{B})$ is a zero of some polynomial $F_j(\mathbf{x})$.

Lecture 3—Proof of Theorem 14

3.1 Singular Points

In proving Theorem 14 we shall begin by considering singular points. A singular point of $F(\mathbf{x}) = 0$ satisfies

$$\frac{\partial F(\mathbf{x})}{\partial x_i} = 0, \quad (1 \le i \le n).$$

Not all the forms $\partial F/\partial x_i$ can be identically zero, since F is absolutely irreducible. Moreover, if one of the partial derivatives is non-zero it will have degree $d-1$, so that it cannot be a multiple of F. Thus if we include a non-zero partial derivative amongst the forms F_j, all singular points will be taken care of.

Our proof of Theorem 14 will use an auxiliary prime p, and we shall also need to account for points which are singular modulo p. We set

$$S(F;B,p) = \{\mathbf{x} \in \mathbb{Z}^n : F(\mathbf{x}) = 0,\ |x_i| \leq B,\ (1 \leq i \leq n),\ p \nmid \nabla F(\mathbf{x})\},$$

and

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

The following lemma then holds.

Lemma 1 *Let $B \geq 2$ and $r = [\log(||F||B)]$, and assume that*

$$P \geq \log^2(||F||B).$$

Then we can find primes $p_1 < \ldots < p_r$ in the range $P \ll p_i \ll P$, such that

$$S(F;B) = \bigcup_{i=1}^{r} S(F;B,p_i).$$

In fact this is the only place in the argument where a dependence on $||F||$ occurs.

For the proof we just pick the first r primes $p_i > AP$, with a suitable constant A. We will then have $P \ll p_i \ll P$ since $P \gg r^2$. For any $\mathbf{x} \in S(F;B)$ there will be some partial derivative $\partial F/\partial x_j$, say, which is non-zero. Using the bound

$$\frac{\partial F}{\partial x_j} \ll_n ||F||B^{d-1},$$

we see that

$$\#\{p > AP : p|\frac{\partial F}{\partial x_j}\} \ll_{n,d} \frac{\log(||F||B)}{\log(AP)}.$$

Thus there are at most $r-1$ such primes, if A is large enough. It follows that one of the primes p_i does not divide $\partial F/\partial x_j$, in which case $\mathbf{x} \in S(F;B,p_i)$, as required.

As a consequence of Lemma 1 it suffices to examine points which are non-singular modulo a fixed prime $p \gg \log^2(||F||B)$, providing that we allow an extra factor $\log(||F||B)$ in our final estimate for k.

3.2 The Implicit Function Theorem

We shall write

$$S(F;B,p) = \bigcup_{\mathbf{t}} S(\mathbf{t}),$$

where

$$S(\mathbf{t}) = \{\mathbf{x} \in S(F;B,p) : \mathbf{x} \equiv \rho\mathbf{t} \pmod{p} \text{ for some } \rho \in \mathbb{Z}\},$$

and $\mathbf{t}$ runs over a set of projective representatives for the non-singular points of $F(\mathbf{t}) = 0$ over $\mathbb{F}_p$.

The proof of Theorem 14 will show that if p is sufficiently large compared with B, then all points $\mathbf{x} \in S(\mathbf{t})$ satisfy an equation $F(\mathbf{x};\mathbf{t}) = 0$. The forms $F(\mathbf{x};\mathbf{t})$ will turn out to have the properties described in Theorem 14, so that if

we take k' to be the number of non-singular points of $F(\mathbf{t}) = 0$ over $\mathbb{F}_p$, we will have $k \ll 1 + k'(\log ||F||B)$, according to the argument above. Indeed we will have $k' \ll p^{n-2} \ll P^{n-2}$, so that it will be enough to show that

$$P \gg B^{(n-1)(n-2)^{-1}} d^{-1/(n-2)} V^{\varepsilon} \log^2 ||F||, \tag{3.1}$$

suffices. Note in particular that (3.1) certainly ensures that $P \gg \log^2(||F||B)$, when B is large enough, so that Lemma 1 applies.

We shall now fix our attention on a particular value of $\mathbf{t}$. Without loss of generality we may take $t_1 = 1$, since we are working in projective space. If the partial derivatives

$$\frac{\partial F}{\partial x_i}(\mathbf{t}) \tag{3.2}$$

were to vanish for $2 \leq i \leq n$, then the first partial derivative must vanish too, since

$$0 = dF(\mathbf{t}) = \mathbf{t}.\nabla F(\mathbf{t}).$$

However $\mathbf{t}$ was assumed to be non-singular, so there must be some non-vanishing partial derivative with $2 \leq i \leq n$. Without loss of generality we shall assume that in fact

$$\frac{\partial F}{\partial x_2}(\mathbf{t}) \neq 0. \tag{3.3}$$

Using Hensel's lemma, along with (3.3), we can lift $\mathbf{t}$ to a p-adic solution $\mathbf{u} \in \mathbb{Z}_p^n$ of $F(\mathbf{u}) = 0$ in which $u_1 = 1$. One can now show that the equation

$$F(1, u_2 + Y_2, u_3 + Y_3, \ldots, u_n + Y_n) = 0$$

may be used to define Y_2 implicitly as a convergent p-adic power series in $Y_3, \ldots, Y_n$, providing that $p|Y_i$ for $2 \leq i \leq n$. This is, in effect, an application of the implicit function theorem, but we shall formulate it in terms of polynomials, as follows.

Lemma 2 *Let $F(\mathbf{x})$ and $\mathbf{u}$ be as above. Then, for any integer $m \geq 1$ we can find $f_m(Y_3, Y_4, \ldots, Y_n) \in \mathbb{Z}_p[Y_3, \ldots, Y_n]$, such that if $F(\mathbf{v}) = 0$ for some $\mathbf{v} \in \mathbb{Z}_p^n$ with $v_1 = 1$ and $\mathbf{v} \equiv \mathbf{u} \pmod p$, then*

$$v_2 \equiv f_m(v_3, \ldots, v_n) \pmod{p^m}. \tag{3.4}$$

We shall prove Lemma 2 by induction on m. Write

$$\frac{\partial F}{\partial x_2}(\mathbf{u}) = \mu,$$

say, and let $f_1(Y_3, \ldots, Y_n) = u_2$, (constant) and

$$f_{m+1}(Y_3, \ldots, Y_n) = f_m(Y_3, \ldots, Y_n) - \mu^{-1} F(1, f_m(Y_3, \ldots, Y_n), Y_3, \ldots, Y_n),$$

for $m \geq 1$. The case $m = 1$ of Lemma 2 is then immediate. For the general case the induction hypothesis yields

$$v_2 \equiv f_m(v_3, \ldots, v_n) \pmod{p^m},$$

so that we may write

$$v_2 = f_m(v_3, \ldots, v_n) + \lambda p^m,$$

with $\lambda \in \mathbb{Z}_p$. Then

$$\begin{aligned} 0 &= F(\mathbf{v}) \\ &\equiv F(1, f_m(v_3, \dots, v_n), v_3, \dots, v_n) \\ &\qquad + \lambda p^m \frac{\partial F}{\partial x_2}(1, f_m(v_3, \dots, v_n), v_3, \dots, v_n) \pmod{p^{m+1}}. \end{aligned} \tag{3.5}$$

Moreover, the induction hypothesis (3.4) shows that

$$f_m(v_3, \dots, v_n) \equiv u_2 \pmod{p},$$

since $\mathbf{v} \equiv \mathbf{u} \pmod{p}$. It follows that

$$\frac{\partial F}{\partial x_2}(1, f_m(v_3, \dots, v_n), v_3, \dots, v_n) \equiv \mu \pmod{p}.$$

We now see from (3.5) that

$$\lambda p^m \equiv -\mu^{-1} F(1, f_m(v_3, \dots, v_n), v_3, \dots, v_n) \pmod{p^{m+1}},$$

so that

$$v_2 \equiv f_{m+1}(v_3, \dots, v_n) \pmod{p^{m+1}}.$$

This completes the induction.

3.3 Vanishing Determinants of Monomials

Clearly we can apply an invertible integral linear transformation to the form $F(\mathbf{x})$ so as to produce a form in which the coefficient of x_n^d is non-zero. Indeed we can find such a transformation in which the coefficients are all $O_{d,n}(1)$. Thus there is no loss of generality in assuming that the monomial x_n^d has non-zero coefficient in $F(\mathbf{x})$.

We now choose a large integer D and define the set

$$\mathcal{E} = \{(e_1, \dots, e_n) \in \mathbb{Z}^n : e_i \geq 0, (1 \leq i \leq n), e_n < d, \sum_{i=1}^{n} e_i = D\}. \tag{3.6}$$

We shall write

$$E = \#\mathcal{E} = \binom{D+n-1}{n-1} - \binom{D-d+n-1}{n-1}, \tag{3.7}$$

and we shall suppose for the moment that $E \leq \#S(\mathbf{t})$. Let $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(E)}$ be distinct vectors in $S(\mathbf{t})$ and let

$$\Delta = \det(\mathbf{x}^{(i)\mathbf{e}})_{1 \leq i \leq E,\, \mathbf{e} \in \mathcal{E}},$$

where we write

$$w_1^{e_1} \dots w_n^{e_n} = \mathbf{w}^{\mathbf{e}}.$$

Thus Δ is an $E \times E$ determinant with rows corresponding to the different vectors $\mathbf{x}^{(i)}$ and columns corresponding to the various exponent n-tuples $\mathbf{e}$.

We proceed to show that Δ is divisible by a large power p^m of p. This will enable us to deduce that Δ vanishes. Since $\mathbf{x}^{(i)} \in S(\mathbf{t})$, we see that the reduction

modulo p of $\mathbf{x}^{(i)}$ represents the same projective point as $\mathbf{t}$ does. It follows that $p \nmid x_1$ so that we may regard $x_1^{-1}\mathbf{x} = \mathbf{v}$, say, as a vector in $\mathbb{Z}_p^n$. We now have $v_1 = t_1 = 1$. Moreover we may lift $\mathbf{t}$ to a vector $(1, u_2, u_3, \dots, u_n) \in \mathbb{Z}_p^n$ on $F = 0$, as in Lemma 2, so that $v_i = u_i + y_i$ for $2 \le i \le n$, for suitable $y_i \in p\mathbb{Z}_p$. Lemma 2 now shows that

$$\Delta = \Big(\prod_{1 \le i \le E} x_1^{(i)} \Big)^D \det(\mathbf{v}^{(i)\mathbf{e}})_{1 \le i \le E,\ \mathbf{e} \in \mathcal{E}} \equiv \Big(\prod_{1 \le i \le E} x_1^{(i)} \Big)^D \Delta_0 \pmod{p^m},$$

where

$$\Delta_0 = \det(M_0), \quad M_0 = (\mathbf{w}^{(i)\mathbf{e}})_{1 \le i \le E,\ \mathbf{e} \in \mathcal{E}},$$

with

$$w_1^{(i)} = 1, \quad w_2^{(i)} = f_m(v_3^{(i)}, \dots, v_n^{(i)}),$$

and

$$w_j^{(i)} = v_j^{(i)} \quad (3 \le j \le n).$$

We proceed to replace $v_j^{(i)}$ by $u_j + y_j^{(i)}$ for $3 \le j \le n$, so that we have $p | y_j^{(i)}$. It follows that

$$\mathbf{w}^{(i)\mathbf{e}} = w_1^{(i)e_1} \dots w_n^{(i)e_n} = g_{\mathbf{e}}(y_3^{(i)}, y_4^{(i)}, \dots, y_n^{(i)})$$

for a suitable collection of polynomials $g_{\mathbf{e}}(Y_3, \dots, Y_n) \in \mathbb{Z}_p[Y_3, \dots, Y_n]$. We now choose an ordering $\prec$ on the vectors

$$\mathbf{f} = (f_3, \dots, f_n), \quad (f_j \in \mathbb{Z}, \ \ f_j \ge 0),$$

in such a way that $\mathbf{f} \prec \mathbf{f}'$ if $\sum f_j < \sum f_j'$. (When $n \ge 4$ this can be done in many different ways.) We then order the monomials $\mathbf{Y}^{\mathbf{f}}$ in the corresponding fashion.

We now perform column operations on M_0 using the following procedure. We take the 'smallest' monomial $\mathbf{Y}^{\mathbf{f}}$, say, occurring in any of the polynomials $g_{\mathbf{e}}$. If this monomial occurs in two or more such polynomials, we use the monomial for which the p-adic order of the coefficient is least. We then interchange columns so as to bring this term into the leading column, and proceed to subtract p-adic integer multiples of the new first column from any other columns which contain the monomial $\mathbf{Y}^{\mathbf{f}}$. Thus this monomial will now occur only in the first column. We then repeat the procedure with the remaining $n-1$ columns, looking again for the 'smallest' monomial, placing it in the second column, and removing it from all later columns. Continuing in this manner we reach an expression

$$\Delta_0 = \det(M_1), \quad M_1 = \big(h_e(y_3^{(i)}, \dots, y_n^{(i)})\big)_{1 \le i \le E,\ 1 \le e \le E},$$

in which we have polynomials $h_e(\mathbf{Y}) \in \mathbb{Z}_p[\mathbf{Y}]$, with successively larger 'smallest' monomial terms.

There are

$$\binom{f+n-3}{n-3} = n(f),$$

say, monomials of total degree f. Hence if $e > n(0) + n(1) + \dots + n(f-1)$, the 'smallest' monomial in $h_e(\mathbf{Y})$ will have total degree at least f. We now recall

that $p|y_j^{(i)}$ for $3 \le j \le n$ whence every element in the e-th column of M_1 will be divisible by p^f. Since

$$\sum_{i=0}^{f} n(i) = \begin{pmatrix} f+n-2 \\ n-2 \end{pmatrix},$$

and

$$\sum_{i=0}^{f} in(i) = (f+1)\begin{pmatrix} f+n-2 \\ n-2 \end{pmatrix} - \begin{pmatrix} f+n-1 \\ n-1 \end{pmatrix},$$

it follows that if

$$\begin{pmatrix} f+n-2 \\ n-2 \end{pmatrix} \le E < \begin{pmatrix} (f+1)+n-2 \\ n-2 \end{pmatrix}, \tag{3.8}$$

then Δ_0 is divisible by

$$p^{n(1)+2n(2)+\ldots+fn(f)+(f+1)(E-n(0)-n(1)-\ldots-n(f))} = p^{\nu},$$

say, where

$$\nu = (f+1)E - \begin{pmatrix} f+n-1 \\ n-1 \end{pmatrix}. \tag{3.9}$$

We therefore specify that the prime power p^m with which we work will have $m = \nu$. This leads to the following conclusion.

Lemma 3 *If E lies in the interval (3.8) and ν is as in (3.9), then*

$$\nu_p(\Delta) \ge \nu.$$

To show that Δ must in fact vanish we shall use information on its size. Each entry in Δ has modulus at most B^D, whence

$$|\Delta| \le E^E B^{DE}.$$

Hence if we impose the condition

$$p^{\nu} > E^E B^{DE}, \tag{3.10}$$

we deduce from Lemma 3 that $\Delta = 0$.

3.4 Completion of the Proof

We are now ready to construct the form F_j corresponding to our chosen value of $\mathbf{t}$. Recall that we have supposed that $\#S(\mathbf{t}) \ge E$, and that we took the points $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(E)}$ to be distinct elements of $S(\mathbf{t})$. We now set $\#S(\mathbf{t}) = K$ and examine the matrix

$$M_2 = \left(\mathbf{x}^{(i)\mathbf{e}}\right)_{1 \le i \le K,\, \mathbf{e} \in \mathcal{E}}$$

where the vectors $\mathbf{x}^{(i)}$ now run over all elements of $S(\mathbf{t})$. From what we have proved it follows that M_2 has rank at most $E-1$. This is trivial if $K \le E-1$, and otherwise the work of the previous section shows that every $E \times E$ minor

vanishes. We therefore see that $M_2\mathbf{c} = \mathbf{0}$ for some non-zero vector $\mathbf{c} \in \mathbb{Z}^E$. Hence if

$$F_j(\mathbf{x}) = \sum_{\mathbf{e}\in\mathcal{E}} c_{\mathbf{e}}\mathbf{x}^{\mathbf{e}}, \tag{3.11}$$

we will have a non-zero form, of degree D, which vanishes for every $\mathbf{x} \in S(\mathbf{t})$. Theorem 14 requires that $F(\mathbf{x})$ does not divide $F_j(\mathbf{x})$. However this is clear from our choice of the exponent set $\mathcal{E}$, since F contains a term in x_n^d, whereas F_j does not.

From (3.10) we see that it suffices to have $p \gg_D B^{DE/\nu}$, so that it will be enough to prove that

$$\lim_{D\to\infty} \frac{DE}{\nu} = \frac{n-1}{n-2} d^{-1/(n-2)}.$$

In view of (3.1) this will complete the proof of Theorem 14.

From (3.7) we have

$$E = \frac{dD^{n-2}}{(n-2)!} + O(D^{n-3}),$$

where the implied constant may depend on n and d. On the other hand, (3.8) implies that

$$E = \frac{f^{n-2}}{(n-2)!} + O(f^{n-3}).$$

We therefore deduce that

$$f = d^{1/(n-2)}D + O(1),$$

whence (3.9) yields

$$\begin{aligned} \nu &= \frac{(n-2)f^{n-1}}{(n-1)!} + O(f^{n-2}) \\ &= d^{(n-1)/(n-2)}(n-2)\frac{D^{n-1}}{(n-1)!} + O(D^{n-2}). \end{aligned} \tag{3.12}$$

These estimates then produce the required limiting behaviour for DE/ν, thereby completing the proof of Theorem 14.

We end this lecture by establishing Theorem 5. For convenience we set $M = (d+1)(d+2)/2$ and $N = d^2+1$. We then suppose that $F(\mathbf{x}) = 0$ has solutions $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)} \in \mathbb{Z}^3$, with $|\mathbf{x}^{(i)}| \ll B$. We now consider a matrix C of size $N \times M$, in which the i-th row consists of the M possible monomials of degree d in the variables $x_1^{(i)}, x_2^{(i)}, x_3^{(i)}$. Let $\mathbf{f} \in \mathbb{Z}^M$ have entries which are the corresponding coefficients of F, so that $C\mathbf{f} = \mathbf{0}$. It follows that C has rank at most $M-1$, since $\mathbf{f}$ is non-zero. We therefore see that $C\mathbf{g} = \mathbf{0}$ has a non-zero integer solution $\mathbf{g}$, which can be constructed from the various sub-determinants of C. It follows that $|\mathbf{g}| \ll_d B^{dM}$. Now take $G(\mathbf{x})$ to be the ternary form, of degree d, corresponding to the coefficient vector $\mathbf{g}$. By our construction, $G(\mathbf{x})$ and $F(\mathbf{x})$ have common zeros at each of the points $\mathbf{x}^{(i)}$ for $1 \le i \le d^2+1$. This will contradict Bézout's Theorem, unless F and G are proportional. In the latter case we may deduce that $||F|| \ll_d ||G|| \ll_d B^{dM}$, since the coefficients of F have no common factor. This gives us the required conclusion.

Lecture 4—Rational Points on Projective Surfaces

This lecture will be devoted to the proofs of Theorems 6 and 8. We shall not present all the details, for which the reader should consult [15].

4.1 Theorem 6—Plane Sections

The principal tool for the proof of Theorem 6 is the following result from the geometry of numbers, for which see [15, Lemma 1, parts (iii) & (iv)].

Lemma 4 *Let $\mathbf{x} \in \mathbb{Z}^n$ lie in the cube $|x_i| \leq B$. Then there is a primitive vector $\mathbf{z} \in \mathbb{Z}^n$, for which $\mathbf{x}.\mathbf{z} = 0$, and such that $|\mathbf{z}| \ll B^{1/(n-1)}$.*

Moreover, if $\mathbf{z} \in \mathbb{Z}^n$ is primitive, then there exist primitive vectors

$$\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(n-1)} \in \mathbb{Z}^n$$

such that

$$|\mathbf{z}| \ll \prod_{j=1}^{n-1} |\mathbf{b}^{(j)}| \ll |\mathbf{z}|. \tag{4.1}$$

These have the property that any vector $\mathbf{x} \in \mathbb{Z}^n$ with $\mathbf{x}.\mathbf{z} = 0$ may be written as a linear combination

$$\mathbf{x} = \lambda_1 \mathbf{b}^{(1)} + \ldots + \lambda_{n-1} \mathbf{b}^{(n-1)} \tag{4.2}$$

with $\lambda_1, \ldots, \lambda_{n-1} \in \mathbb{Z}$ and

$$\lambda_j \ll \frac{|\mathbf{x}|}{|\mathbf{b}^{(j)}|}, \quad (1 \leq j \leq n-1).$$

It follows that the region $|\mathbf{x}| \leq X$ contains $O(X^{n-1}/|\mathbf{z}|)$ integral vectors orthogonal to $\mathbf{z}$, providing that $X \gg |\mathbf{z}|$.

For the last part we note that $|\mathbf{b}^{(j)}| \ll |\mathbf{z}| \ll X$, by (4.1), and hence that there are $O(X/|\mathbf{b}^{(j)}|)$ choices for each λ_j. The result then follows from a second application of (4.1).

Taking $n = 4$ we see that every relevant point on the surface $F(\mathbf{x}) = 0$ must lie on one of $O(B^{4/3})$ planes $\mathbf{x}.\mathbf{y} = 0$ with $|\mathbf{y}| \ll B^{1/3}$. We proceed to count the number of points of the surface $F(\mathbf{x}) = 0$ which lie on a given plane. According to Lemma 4, each vector $\mathbf{y}$ determines a triple of vectors $\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \mathbf{b}^{(3)}$. We may then write $\mathbf{x}$ in the form (4.2) and substitute into the equation $F(\mathbf{x}) = 0$ to obtain a condition $G(\lambda_1, \lambda_2, \lambda_3) = 0$, say, in which G is an integral form of degree d, though not necessarily irreducible. If $\mathbf{x}$ is primitive then it is clear that $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{Z}^3$ is also primitive. Moreover, if we choose $|\mathbf{b}^{(1)}| \leq |\mathbf{b}^{(2)}| \leq |\mathbf{b}^{(3)}|$, then the condition $\max |x_i| \leq B$ implies $\max |\lambda_i| \leq cB/|\mathbf{b}^{(1)}|$ for some absolute constant c.

4.2 Theorem 6—Curves of Degree 3 or more

If H is an absolutely irreducible factor of G, then the solutions of $H(\lambda_1, \lambda_2, \lambda_3) = 0$ will correspond to points on an irreducible plane curve in the surface $F = 0$. If H has degree at least 3, then

$$N(H; cB/|\mathbf{b}^{(1)}|) \ll_\varepsilon (\frac{B}{|\mathbf{b}^{(1)}|})^{2/3+\varepsilon}, \tag{4.3}$$

by Theorem 2. It is important to notice that the implied constant is independent of H, and hence of the vectors $\mathbf{b}^{(j)}$ and $\mathbf{y}$. The estimate (4.3) allows us to bound the number of solutions of $G(\lambda_1, \lambda_2, \lambda_3) = 0$ arising from all factors H of G with degree 3 or more. We proceed to sum the bound (4.3) over those vectors $\mathbf{y}$ that arise. In order to do this we need to count how many vectors $\mathbf{y}$ can correspond to a given $\mathbf{b}^{(1)}$. To do this we apply the final part of Lemma 4, taking $\mathbf{z}$ to be $\mathbf{b}^{(1)}$. We then see that there are $O(B/|\mathbf{b}^{(1)}|)$ possible vectors $\mathbf{y}$ in the region $|\mathbf{y}| \ll B^{1/3}$ which are orthogonal to a given $\mathbf{b}^{(1)}$, so that the total contribution of the estimates (4.3), when we sum over $\mathbf{y}$, is

$$\ll_\varepsilon \sum (\frac{B}{|\mathbf{b}^{(1)}|})^{5/3+\varepsilon}, \tag{4.4}$$

the sum being over the possible vectors $\mathbf{b}^{(1)}$. Since we took $|\mathbf{b}^{(1)}| \le |\mathbf{b}^{(2)}| \le |\mathbf{b}^{(3)}|$, and we also have

$$|\mathbf{y}| \ll \prod_{j=1}^{n-1} |\mathbf{b}^{(j)}| \ll |\mathbf{y}|$$

by (4.1), it follows that $|\mathbf{b}^{(1)}| \ll |\mathbf{y}|^{1/3} \ll B^{1/9}$. The sum (4.4) is therefore

$$\ll_\varepsilon B^{5/3+\varepsilon}(B^{1/9})^{4-5/3-\varepsilon} \ll_\varepsilon B^{52/27+\varepsilon}.$$

This is satisfactory for Theorem 6.

4.3 Theorem 6—Quadratic Curves

It remains to handle the case in which $(\lambda_1, \lambda_2, \lambda_3)$ is a zero of a linear or quadratic factor H of G. Here we shall be brief. If H is linear, then the equations $\mathbf{x}.\mathbf{y} = 0$ and $H(\lambda_1, \lambda_2, \lambda_3) = 0$ describe a line in the surface $F = 0$, so that the corresponding points $\mathbf{x}$ are not counted by $N_1(F; B)$. According to Harris [10, Proposition 18.10], the generic plane section of any irreducible hypersurface is itself irreducible. Thus the set of vectors $\mathbf{y}$ for which the corresponding form G is reducible must lie on a certain union of irreducible surfaces in $\mathbb{P}^3$ (or possibly indeed in some smaller algebraic set). One can show [15, §6] that the number and degrees of the components of this set may be bounded in terms of d, so that there are $O(Y^3)$ admissible vectors $\mathbf{y}$ with $|\mathbf{y}| \le Y$. To obtain the estimate $O(Y^3)$ one may apply Pila's result (2.3) to components of degree 2 or more, and the trivial bound for linear components. In the case in which H is quadratic it can be shown that there are $O_\varepsilon(B^{1+\varepsilon}|\mathbf{y}|^{-1/3})$ solutions $(\lambda_1, \lambda_2, \lambda_3)$. Thus vectors $\mathbf{y}$ with $Y/2 < |\mathbf{y}| \le Y$ contribute a total $O_\varepsilon(Y^3 B^{1+\varepsilon} Y^{-1/3})$ to $N_1(F; B)$. If we sum over dyadic intervals with $Y \ll B^{1/3}$ the result is a contribution $O(B^{17/9+\varepsilon})$. This is clearly satisfactory for Theorem 6. The reader may note that it is only for cubic surfaces that the exponent $52/27$ is required. In all other cases one can do better.

4.4 Theorem 8—Large Solutions

We turn now to Theorem 8. For the remainder of this lecture, all implied constants may depend on the form G. This dependence will not be mentioned explicitly. It will be convenient to make a change of variable in G so that the coefficient of x^d is positive. This will not affect the result at all.

The reader may care to note that our treatment differs somewhat from that presented in [15, §7]. This is because we have made the simplifying assumption that the binary form G is irreducible over $\mathbb{Q}$.

One technical difficulty with the proof of Theorem 8 is that one may have a value of $G(x,y)$ in the range $1 \le G(x,y) \le X$, for which x, y are considerably larger than $X^{1/d}$. The first task is to show that this happens relatively rarely. We assume that $C \gg X^{1/d}$ and define

$$S(X,C) = \\ \#\{(x,y) \in \mathbb{Z}^2 : 1 \le G(x,y) \le X,\ C < \max(|x|,|y|) \le 2C,\ \text{h.c.f.}(x,y) = 1\},$$

Now, if x, y is counted by $S(X,C)$ then there is at least one factor $x - ay$ of $G(x,y)$ for which $|x - ay| \ll X^{1/d}$. Hence if we take $C \ge cX^{1/d}$ with a sufficiently large constant c, we must have

$$C \ll |x - a'y| \ll C$$

for every other factor $x - a'y$ of $G(x,y)$. It then follows that

$$|x - ay| \ll XC^{1-d}. \tag{4.5}$$

Since Roth's Theorem implies that $|x - ay| \gg_\varepsilon C^{-1-\varepsilon}$ we deduce that $C \ll X^2$. (In fact we may draw a stronger conclusion, but the above suffices.) Thus we will have $S(X,C) = 0$ unless $C \ll X^2$.

We now estimate the contribution to $S(X,C)$ arising from pairs (x,y) for which (4.5) holds with a particular value of a. Such pairs produce primitive lattice points in the parallelogram $|y| \le 2C,\ |x - ay| \ll XC^{1-d}$. In general a parallelogram of area A, centred on the origin, will contain $O(1 + A)$ primitive lattice points (see [15, Lemma 1, part (vii)]). We therefore deduce that

$$S(X,C) \ll 1 + XC^{2-d}.$$

We proceed to sum this up, for dyadic ranges with $C \ll X^2$. Thus, if

$$S'(X,C) = \\ \#\{(x,y) \in \mathbb{Z}^2 : 1 \le G(x,y) \le X,\ \max(|x|,|y|) > C,\ \text{h.c.f.}(x,y) = 1\}$$

we will deduce that

$$S'(X,C) \ll \log X + XC^{2-d},$$

when $C \gg X^{1/d}$.

We proceed to define

$$r(n) = \#\{(x,y) \in \mathbb{Z}^2 : n = G(x,y)\}$$

and

$$r_1(n;C) = \#\{(x,y) \in \mathbb{Z}^2 : n = G(x,y),\ \max(|x|,|y|) \le C\},$$

$$r_2(n;C) = \#\{(x,y) \in \mathbb{Z}^2 : n = G(x,y),\ \max(|x|,|y|) > C\}.$$

Thus

$$\begin{aligned}\sum_{n\le X} r_2(n;C) &= \sum_{h\ll X^{1/d}} S'(\frac{X}{h^d}, \frac{C}{h}) \\ &\ll \sum_{h\ll X^{1/d}} \{\log X + \frac{X}{h^d}(\frac{C}{h})^{2-d}\} \\ &\ll X^{1/d}\log X + XC^{2-d}\sum_{h\ll X^{1/d}} h^{-2} \\ &\ll X^{1/d}\log X + XC^{2-d}. \end{aligned} \tag{4.6}$$

This shows that 'large' pairs (x,y) make a relatively small contribution in our problem.

It is trivial that

$$\sum r_1(n;C) \ll C^2, \tag{4.7}$$

whence

$$\sum_{n\le X} r(n) \ll C^2 + X^{1/d}\log X + XC^{2-d} \ll X^{2/d},$$

providing that we choose $C = cX^{1/d}$ with a suitable constant c. It follows in particular that there are $O(X^{2/d})$ positive integers $n \le X$ represented by G.

4.5 Theorem 8—Inequivalent Representations

For any n which has two inequivalent representations by the form $G(x,y)$, we must either have $r_2(n;C) > 0$, or we will produce a point (x_1,x_2,x_3,x_4) on the surface

$$E(\mathbf{x}) = G(x_1,x_2) - G(x_3,x_4) = 0,$$

in the cube $|x_i| \le C$, and such that (x_1,x_2) and (x_3,x_4) are not related by an automorphism. We define $\mathcal{N}(C)$ to be the number of such points. We shall show that

$$\mathcal{N}(C) \ll_\varepsilon C^{52/27+\varepsilon}, \tag{4.8}$$

and that the form G has $O(1)$ automorphisms. However before proving this, we show how Theorem 8 will follow.

We first observe that (4.7) and (4.8) imply the estimate

$$\sum_{n\le X} r_1(n;C)^2 \le \mathcal{N}(C) + O(\sum_{n\le X} r_1(n;C)) \ll C^2,$$

where the second sum counts pairs (x_1,x_2) and (x_3,x_4) which are related by one of the finitely many automorphisms. If $C = cX^{1/d}$ with a sufficiently small constant c, then $\max(|x|,|y|) \le C$ implies $|G(x,y)| \le X$. Since we are assuming that $G(1,0) > 0$, it is then trivial that

$$\sum_{n\le X} r_1(n;C) \gg C^2$$

since a positive proportion of the pairs x, y with $\max(|x|, |y|) \le C$ will have $G(x, y) > 0$. Now Cauchy's inequality yields

$$\sum_{n \le X, r_1(n:C)>0} 1 \ge \frac{\{\sum_{n \le X} r_1(n; C)\}^2}{\sum_{n \le X} r_1(n; C)^2} \gg C^2.$$

We therefore see that the number of positive integers $n \le X$ which are represented by G, has exact order $X^{2/d}$, as required for Theorem 8.

It remains to give a non-trivial bound for the number of integers n with two or more essentially different representations by G. Since such integers are either counted by $\mathcal{N}$ or have $r_2(n; C) > 0$, we see that the number of them is

$$\begin{aligned} &\le \quad \mathcal{N}(C) + \sum_{n \le X} r_2(n; C) \\ &\ll_\varepsilon \quad C^{52/27+\varepsilon} + X^{1/d} \log X + XC^{2-d}, \end{aligned}$$

by (4.6) and (4.8). We therefore obtain a bound $O_\varepsilon(X^{52/(27d-2)+\varepsilon})$ on taking $C = X^{27/(27d-2)}$. This proves Theorem 8, subject to the claims made above.

4.6 Theorem 8—Points on the Surface $E = 0$

It remains to investigate integral points on the surface

$$E(\mathbf{x}) = G(x_1, x_2) - G(x_3, x_4) = 0,$$

for which $\max |x_i| \le C$. Using the fact that G is irreducible one may show (see [15, §7]) that E has no rational linear or quadratic factor. Thus one may apply Theorem 6 to each factor of $E(\mathbf{x})$ to deduce that

$$N_1(E; Y) \ll_\varepsilon Y^{52/27+\varepsilon}.$$

We now write $\mathcal{N}^{(*)}(C)$ to denote the number of integral zeros of E, not necessarily primitive, lying in the cube $|x_i| \le C$, but not on any line in the surface $E = 0$. Then

$$\begin{aligned} \mathcal{N}^{(*)}(C) \quad &= \quad 1 + \sum_{h \ll C} N_1(E; B/h) \\ &\ll_\varepsilon \quad 1 + \sum_h (C/h)^{52/27+\varepsilon} \\ &\ll_\varepsilon \quad C^{52/27+\varepsilon}. \end{aligned}$$

Points which do not lie on lines in the surface $E = 0$ therefore make a contribution which is satisfactory for (4.8).

Since G has no repeated factors, the surface E is non-singular. Thus Colliot-Thélène's result, Theorem 12, shows that E contains finitely many lines L, say. Now if L is not defined over $\mathbb{Q}$ it can have at most $O(C)$ integral points in the cube $\max |x_i| \le C$. Such lines therefore make a satisfactory total contribution $O(C)$ to (4.8).

For any line L lying in the surface $E(\mathbf{x}) = 0$ one can show that either all points on L satisfy $G(x_1, x_2) = G(x_3, x_4) = 0$, or that the points on L may be written as $(x, y, a_1x + a_2y, a_3x + a_4y)$ with $a_1a_4 \ne a_2a_3$, so that one has

$$G(a_1x + a_2y, a_3x + a_4y) = G(x, y) \tag{4.9}$$

identically. In the first case the points on L correspond to solutions of $G(x,y) = n$ with $n = 0$, which is excluded. In the second case we produce an automorphism of G, and if L is defined over $\mathbb{Q}$ this will be a rational automorphism. Thus inequivalent solutions of $G(x,y) = n > 0$ cannot lie on rational lines in the surface $E = 0$.

Finally we note that all automorphisms produce lines in the surface, as in (4.9). Thus there can be finitely many such automorphisms.

Lecture 5—Affine Varieties

5.1 Theorem 15—The Exponent Set $\mathcal{E}$

In this lecture we shall prove Theorem 15, which concerns the counting function (2.12) for integer points on affine varieties. We shall also illustrate Theorem 15 by deriving a new result on the representation of k-free numbers by polynomials.

We begin by following the previous proof of Theorem 14. Thus if we take $B = \max(B_1, \dots, B_n)$ then we may use an auxiliary prime p in the range $P \ll p \ll P$, providing that $P \gg \log^2(||F||B)$. We will take $(t_1, \dots, t_n)$ to be a non-singular point on the variety $F(\mathbf{t}) \equiv 0 \pmod p$. We then aim to find a polynomial $F_j(\mathbf{x})$, determined by $\mathbf{t}$, such that $F_j(\mathbf{x}) = 0$ for all integral points $\mathbf{x} \equiv \mathbf{t} \pmod p$ that are counted by $N(F;\mathbf{B})$. This will be possible if P is sufficiently large, and we will then be able to take $k \ll P^{n-1}\log(||F||B)$ in Theorem 15, since there are $O(p^{n-1})$ possible vectors $\mathbf{t}$ modulo p. The proofs of Lemmas 2 and 3 now go through just as before.

One major difference in our new situation lies in the fact that the values of B_i may be of unequal magnitudes. Thus one cannot arrange for F to have a non-zero term in x_n^d, say, merely by using a linear change of variables, since this could radically alter the sizes of the B_i. We therefore use a different choice for the exponent set $\mathcal{E}$. The set $\mathcal{E}$ must firstly allow us still to conclude that $F(\mathbf{x})$ does not divide $F_j(\mathbf{x})$, and secondly permit us to bound the determinant Δ as sharply as possible.

To achieve the first goal, we write

$$F(X_1, \dots, X_n) = \sum_{\mathbf{f}} a_{\mathbf{f}} X_1^{f_1} \dots X_n^{f_n},$$

and let $P(F)$ be the Newton polyhedron of F, defined as the convex hull of the points $\mathbf{f} \in \mathbb{R}^n$ for which $a_{\mathbf{f}} \neq 0$. There may be more than one exponent vector $\mathbf{f}$ for which $\mathbf{B}^{\mathbf{f}}$ takes the maximal value T, but in any case the maximum will be achieved for at least one vertex $\mathbf{m}$, say of P, so that

$$T = \mathbf{B}^{\mathbf{m}}. \tag{5.1}$$

We shall then define our exponent set $\mathcal{E}$ to be

$$\mathcal{E} = \{(e_1, \dots, e_n) \in \mathbb{Z}^n : e_i \geq 0, (1 \leq i \leq n), \sum_{i=1}^{n} e_i \log B_i \leq Y, \\ e_i < m_i \text{ for at least one } i\}. \tag{5.2}$$

We now show that the choice (5.2) ensures that $F(\mathbf{x}) \nmid F_j(\mathbf{x})$. For any two polynomials $G_1, G_2 \in \mathbb{R}[\mathbf{x}]$ we have $P(G_1G_2) = P(G_1) + P(G_2)$, (Ostrowski

[27]). Now, for any polynomial G, the set $P(F) + P(G)$ must contain a point $\mathbf{m} + \mathbf{g}$ for some exponent vector $\mathbf{g}$ of G. Since $\mathbf{g}$ has non-negative components it follows that $\mathbf{m} + \mathbf{g} \notin \mathcal{E}$. Thus we cannot have $F_j = FG$, whence F cannot divide F_j, as required.

5.2 Completion of the Proof of Theorem 15

It remains to estimate the size of the determinant Δ and to compute the exponent ν for our new set $\mathcal{E}$.

Tackling the determinant Δ first, we use the fact that $|x_i^{(j)}| \leq B_i$, to show that the column corresponding to exponent vector $\mathbf{e}$ consists of elements of modulus at most

$$\mathbf{B}^{\mathbf{e}} = \prod_{i=1}^{n} B_i^{e_i}.$$

It follows that

$$|\Delta| \leq E^E \mathbf{B}^{\mathbf{E}}, \tag{5.3}$$

where

$$\sum_{\mathbf{e} \in \mathcal{E}} \mathbf{e} = \mathbf{E}. \tag{5.4}$$

For any $\mathbf{e} \in \mathbb{Z}^n$ with $e_i \geq 0$ we set

$$\sigma(\mathbf{e}) = \sum_{i=1}^{n} e_i \log B_i.$$

Then

$$\log \mathbf{B}^{\mathbf{E}} = \sum_{\mathbf{e} \in \mathcal{E}} \sigma(\mathbf{e}) = \sum_{\sigma(\mathbf{e}) \leq Y} \sigma(\mathbf{e}) - \sum\nolimits^{(1)} \sigma(\mathbf{e}),$$

where $\Sigma^{(1)}$ denotes the conditions $\sigma(\mathbf{e}) \leq Y$ and $e_i \geq m_i$ for $1 \leq i \leq n$. If we substitute $e_i + m_i$ for e_i in $\Sigma^{(1)}$ we obtain

$$\begin{aligned} \log \mathbf{B}^{\mathbf{E}} &= \sum_{\sigma(\mathbf{e}) \leq Y} \sigma(\mathbf{e}) - \sum_{\sigma(\mathbf{e}) \leq Y - \log T} (\sigma(\mathbf{e}) + \log T) \\ &= \sum_{Y - \log T < \sigma(\mathbf{e}) \leq Y} \sigma(\mathbf{e}) - (\log T) \sum_{\sigma(\mathbf{e}) \leq Y - \log T} 1 \\ &= \{Y + O(\log T)\} \sum_{Y - \log T < \sigma(\mathbf{e}) \leq Y} 1 \\ &\qquad -(\log T) \sum_{\sigma(\mathbf{e}) \leq Y - \log T} 1. \end{aligned} \tag{5.5}$$

Now, if $B = \max B_i$, then

$$\#\{\mathbf{e} : \sigma(\mathbf{e}) \leq Z\} = \frac{Z^n}{n! \prod \log B_i} + O\left(\frac{Z^{n-1}}{\prod \log B_i} \log B\right) \tag{5.6}$$

for $Z \geq 0$, as an easy induction on n shows. Thus

$$(\log T) \sum_{\sigma(\mathbf{e}) \leq Y - \log T} 1 = \frac{Y^n}{n! \prod \log B_i} \log T + O\left(\frac{Y^{n-1}}{\prod \log B_i} \log^2 T\right) \tag{5.7}$$

whether $Y \geq \log T$ or not.

It is not so easy to estimate the first term in (5.5). However for any $\delta \in (0,1]$ we have

$$\int_Z^{Z(1+\delta)} \{ \sum_{Y-\log T < \sigma(\mathbf{e}) \leq Y} 1\} dY = \sum_{\mathbf{e}} \int 1 \, dY,$$

with the range of integration on the right being for $Z \leq Y \leq Z(1+\delta)$ and $\sigma(\mathbf{e}) \leq Y \leq \sigma(\mathbf{e}) + \log T$. Thus (5.6) yields

$$\begin{aligned} \int_Z^{Z(1+\delta)} \{ \sum_{Y-\log T < \sigma(\mathbf{e}) \leq Y} 1\} dY &\leq (\log T) \sum_{Z-\log T \leq \sigma(\mathbf{e}) \leq Z(1+\delta)} 1 \\ &= \frac{\log T}{\prod \log B_i} \{ \frac{(1+\delta)^n - 1}{n!} Z^n + O(Z^{n-1} \log T)\}. \end{aligned}$$

Similarly we find that

$$\begin{aligned} \int_Z^{Z(1+\delta)} \{ \sum_{Y-\log T < \sigma(\mathbf{e}) \leq Y} 1\} dY &\geq (\log T) \sum_{Z \leq \sigma(\mathbf{e}) \leq Z(1+\delta)-\log T} 1 \\ &= \frac{\log T}{\prod \log B_i} \{ \frac{(1+\delta)^n - 1}{n!} Z^n + O(Z^{n-1} \log T)\}. \end{aligned}$$

Thus there is some Y in the range $Z \leq Y \leq Z(1+\delta)$ for which

$$\begin{aligned} \sum_{Y-\log T < \sigma(\mathbf{e}) \leq Y} 1 &= \frac{\log T}{\prod \log B_i} \{ \frac{(1+\delta)^n - 1}{\delta n!} Z^{n-1} + O(\delta^{-1} Z^{n-2} \log T)\} \\ &= \frac{\log T}{\prod \log B_i} \{ \frac{Z^{n-1}}{(n-1)!} + O(\delta Z^{n-1}) + O(\delta^{-1} Z^{n-2} \log T)\} \\ &= \frac{\log T}{\prod \log B_i} \{ \frac{Y^{n-1}}{(n-1)!} + O(\delta Y^{n-1}) + O(\delta^{-1} Y^{n-2} \log T)\}. \end{aligned}$$

If $Y \geq \log T$ we may choose

$$\delta = \sqrt{\frac{\log T}{Y}}$$

so that

$$\sum_{Y-\log T < \sigma(\mathbf{e}) \leq Y} 1 = \frac{\log T}{\prod \log B_i} \{ \frac{Y^{n-1}}{(n-1)!} + O(Y^{n-3/2} (\log T)^{1/2})\}. \tag{5.8}$$

If $Z \geq \log T$, we now see that any range $Z \leq Y \leq 2Z$ contains a value of Y for which (5.8) holds. For such Y equations (5.5), (5.7) and (5.8) yield

$$\begin{aligned} \log \mathbf{B}^{\mathbf{E}} &= \frac{\log T}{\prod \log B_i} \frac{Y^n}{(n-1)!} - \frac{\log T}{\prod \log B_i} \frac{Y^n}{n!} + O(\frac{\log T}{\prod \log B_i} Y^{n-1/2} (\log T)^{1/2}) \\ &= \frac{n-1}{n!} \frac{\log T}{\prod \log B_i} Y^n \{1 + O(\sqrt{\frac{\log T}{Y}})\}. \end{aligned} \tag{5.9}$$

A similar but simpler argument provides us with our estimate for ν. Indeed (3.8) and (3.9) still apply, but with n replaced by $n+1$. It follows that

$$\nu = \frac{1}{n(n-2)!} \{(n-1)! E\}^{n/(n-1)} \{1 + O(E^{-1/(n-1)})\},$$

where $E = \#\mathcal{E}$. However, as in (5.5), we have

$$\#\mathcal{E} = \sum_{\sigma(\mathbf{e}) \le Y} 1 - \sum_{\sigma(\mathbf{e}) \le Y - \log T} 1 = \sum_{Y - \log T < \sigma(\mathbf{e}) \le Y} 1.$$

For the value of Y we have chosen, it follows from (5.8) that

$$E = \frac{\log T}{\prod \log B_i} \frac{Y^{n-1}}{(n-1)!} \{1 + O(\sqrt{\frac{\log T}{Y}})\}. \tag{5.10}$$

Since $\log B_i \le \log T$ for each index i, we deduce that

$$E \gg (\frac{Y}{\log T})^{n-1}$$

if $Y \gg \log T$, and hence that

$$\nu = \frac{Y^n}{n(n-2)!} \{\frac{\log T}{\prod \log B_i}\}^{n/(n-1)} \{1 + O(\sqrt{\frac{\log T}{Y}})\}. \tag{5.11}$$

As before, we need $p^\nu > |\Delta|$. A comparison of (5.3) (5.9) and (5.11) now shows that if $E \le C$, say, then we can take

$$\log p = \frac{\log \mathbf{B}^{\mathbf{E}}}{\nu} + O_C(1) = \{\frac{\prod \log B_i}{\log T}\}^{1/(n-1)} \{1 + O(\sqrt{\frac{\log T}{Y}})\} + O_C(1).$$

If $\varepsilon > 0$ is such that $B_i \ge B^\varepsilon$ for $1 \le i \le n$, and $Y = \lambda \log T$, then (5.10) yields $E \ll_{\lambda,\varepsilon,d} 1$. By allowing λ to become arbitrarily large we can therefore take

$$\log p \le (1 + \varepsilon) \{\frac{\prod \log B_i}{\log T}\}^{1/(n-1)}.$$

Thus the bound given in Theorem 15 certainly holds if $B_i \ge B^\varepsilon$ for $1 \le i \le n$.

On the other hand, if $B_1 \le B^\varepsilon$, say, we can merely take the polynomials $F_j(\mathbf{x})$ as $F_j(\mathbf{x}) = x_1 - a_j$ for the various integers $a_j \in [-B_1, B_1]$. Clearly each relevant point will be a zero of such a polynomial, and there are $\ll B_1 \ll T^\varepsilon$ of them. This completes the proof of Theorem 15.

5.3 Power-Free Values of Polynomials

We shall now apply Theorem 15 to the problem of power-free values of polynomials. Let $f[X] \in \mathbb{Z}[X]$ be an irreducible polynomial of degree d. It can happen that there is a prime p such that $p^2 | f(n)$ for every integer n, even though f is necessarily primitive. (The polynomial $X^4 + 2X^3 - X^2 - 2X + 4$ provides an example, with $p = 2$.) However if we assume that there is no prime p such that $p^2 | f(n)$ for every integer n, then we would expect $f(n)$ to represent infinitely many square-free integers. This is known only for polynomials of degree $d \le 3$. This is relatively trivial for $d \le 2$, and was shown by Erdős [6] for $d = 3$. When one considers polynomials of higher degree one may ask for which values of k one can assert that $f(n)$ is infinitely often k-th power free, or "k-free" for short. Hooley showed that one may take $k = d - 1$ in general, and Nair [25] that any $k \ge (\sqrt{2} - \frac{1}{2})d$ is admissible.

We shall strengthen this result for $d \ge 10$ as follows.

Theorem 16 *Let $f[X] \in \mathbb{Z}[X]$ be an irreducible polynomial of degree d, with positive leading coefficient. Suppose we have an integer $k \geq (3d+2)/4$, and assume moreover that there is no prime p such that $p^k | f(n)$ for every integer n. Then*

$$\#\{n \leq B : f(n) \text{ is } k\text{-free}\} \sim C(k,f)B$$

as B tends to infinity, where

$$C(k,f) = \prod_p (1 - \rho(p^k)/p^k),$$

with $\rho(n)$ being the number of zeros modulo n of the polynomial $f(x)$.

Note that although $(3d+2)/4 < (\sqrt{2} - \frac{1}{2})d$ as soon as $d \geq 4$, we only have $\lceil (3d+2)/4 \rceil < \lceil (\sqrt{2} - \frac{1}{2}) \rceil d$ when $d \geq 10$.

The initial stages of the argument are straightforward. We shall assume that f is fixed, so that any order constants in the following may depend on f.

One has

$$\sum_{h^k | f(n)} \mu(h) = \begin{cases} 1, & f(n) \text{ is } k\text{-free}, \\ 0, & \text{otherwise}, \end{cases}$$

whence

$$\#\{n \leq B : f(n) \text{ is } k\text{-free}\} = \sum_h \mu(h) N(h,B),$$

with

$$N(h,B) = \#\{n \leq B : h^k | f(n)\}.$$

Moreover we see that $N(h,B) = 0$ for $h \gg B^{d/k}$, and that

$$N(h,B) = \frac{B}{h^k}\rho(h^k) + O(\rho(h^k))$$

in general. Since $\rho(h)$ is multiplicative, and $\rho(p) \ll 1$, we see that $\rho(h^k) \ll_\varepsilon h^\varepsilon$ for any $\varepsilon > 0$. It follows that

$$\begin{aligned} \sum_{h \leq H} \mu(h) N(h,B) &= B \sum_{h \leq H} \mu(h)\rho(h)/h^k + O_\varepsilon(\sum_{h \leq H} h^\varepsilon) \\ &= B\{C(k,f) + O_{f,\varepsilon}(H^{\varepsilon - k})\} + O_{f,\varepsilon}(H^{1+\varepsilon}). \end{aligned}$$

Thus, if we choose $H = B^{1-\delta}$ with any fixed positive δ we will have

$$\sum_{h \leq B^{1-\delta}} \mu(h) N(h,B) \sim C(k,f)B.$$

It therefore remains to show that

$$\sum_{B^{1-\delta} < h \ll B^{d/k}} N(h,B) = o(B). \tag{5.12}$$

For any integer H with $B^{1-\delta} \ll H \ll B^{d/k}$, we have

$$\sum_{H < h \leq 2H} N(h,B) \leq N(F; B_1, B_2, B_3),$$

where $F(x_1, x_2, x_3) = f(x_1) - x_2^k x_3$ and $B_1 = B$, $B_2 = 2H$, $B_3 \ll B^d/H^k$. This polynomial F is clearly absolutely irreducible. It will be convenient to replace the above inequality by a slightly sharper one

$$\sum_{H<h\leq 2H} N(h, B) \leq N'(F; B_1, B_2, B_3), \tag{5.13}$$

in which $N'(F; B_1, B_2, B_3)$ counts only solutions in which x_2 is a positive square-free integer. We now apply Theorem 15, in which T will be of order B^d. Writing

$$\eta = \frac{\log H}{\log B}$$

so that

$$1 - \delta \leq \eta \leq d/k + o(1)$$

we find that all relevant triples satisfy one of

$$O_{d,\varepsilon}(B^{2\sqrt{\eta(1-k\eta/d)}+\varepsilon}) \tag{5.14}$$

auxiliary equations $F_j(x_1, x_2, x_3) = 0$, of degree at most $D = D(d, \varepsilon)$. On multiplying by x_2^{kD} we can eliminate x_3 from such an auxiliary equation, using the fact that

$$x_2^k x_3 = f(x_1). \tag{5.15}$$

It follows that we may assume that the auxiliary equations take the shape $G_j(x_1, x_2) = 0$, with no dependence on x_3. This elimination process will produce produce polynomials G_j which do not vanish identically, since the original polynomials F_j were coprime to F. Moreover the degree of each G_j is at most $D(d, \varepsilon)$ for some new fuction D. If G_j is not absolutely irreducible we may split it into its irreducible factors. In view of Theorem 3 we can concentrate on factors which have rational coefficients. Thus we may suppose that $G_j(x_1, x_2)$ is absolutely irreducible.

We now write $H_j(y_1, y_2, y_3) = y_3^D G_j(y_1/y_3, y_2/y_3)$, so that H_j is a non-zero form of degree D. According to Theorem 5 there are two cases. In the first case $N(H_j; L) \leq D^2$ where $L = \max(B, 2H, 1)$. It then follows that j determines $O_{d,\varepsilon}(1)$ admissible values for x_1, x_2. Moreover, since we are assuming that $x_2 > 0$, the relation (5.15) then shows that any admissible pair x_1, x_2 determines at most one admissible value of x_3. In the second case the coefficients of G_j may be taken to be integers of size $O(B^N)$ for some exponent N depending at most on d and ε. Since we are only interested in values $x_2 > 0$ we may remove any factors of x_2 from the polynomial G_j. Thus we can write

$$G_j(X_1, X_2) = G_j^{(0)}(X_1) + X_2 G_j^{(1)}(X_1, X_2)$$

for appropriate integral polynomials $G_j^{(0)}$ and $G_j^{(1)}$, with $G_j^{(0)}$ not identically zero, and having coefficients $O(B^N)$. We now subdivide this second case into 2 subcases. In the first subcase we will have $f(x) \nmid G_j^{(0)}(x)$, while in the second subcase we will have $f(x) | G_j^{(0)}(x)$

For the first of these subcases we observe that x_2 must divide $G_j^{(0)}(x_1)$. However it is clear from (5.15) that x_2 also divides $f(x_1)$. Thus we conclude

that x_2 must be a factor of the resolvent R of f and $G_j^{(0)}$. Since $f(X)$ does not divide $G_j^{(0)}(X)$ it follows that R is non-zero, and our bound on the coefficients of $G_j^{(0)}$ implies that $R \ll B^{N'}$ for some N' depending only on d and ε. Hence x_2 can take at most $d(R) \ll_\varepsilon B^\varepsilon$ values in this first subcase. (Here $d(\dots)$ is the usual divisor function.) Now x_2 is assumed to be square-free, so that $\rho(x_2) \ll B^\varepsilon$. Moreover the equation (5.15) implies that $f(x_1) \equiv 0 \,(\mathrm{mod}\, x_2)$. Hence we see that the number of possible values of x_1 corresponding to each admissible x_2 is

$$\ll (1 + B/x_2)\rho(x_2) \ll_\varepsilon (1 + B/H)B^\varepsilon \ll_\varepsilon B^{\delta+\varepsilon}.$$

As before x_1 and x_2 determine x_3, by (5.15). It therefore follows that each value of j produces $O(B^{\delta+2\varepsilon})$ triples (x_1, x_2, x_3) in the first subcase.

Turning to the second subcase, in which $F(X)|G_j^{(0)}(X)$, we note that $G_j^{(0)}$ must have degree at least d. We then apply Theorem 15 to the equation $G_j(x_1, x_2) = 0$, taking $B_1 = B$, $B_2 = 2H$, and $T \geq B^d$. This shows that there are $O_\varepsilon(B^\varepsilon H^{1/d})$ possible pairs x_1, x_2, and as usual any such pair determines at most one value of x_3.

We can finally conclude, via (5.14), that (5.13) holds with

$$N'(F; B_1, B_2, B_3) \ll B^{2\sqrt{\eta(1-k\eta/d)}+\eta/d+2\varepsilon},$$

providing that δ is taken to be small enough. This suffices to establish (5.12), assuming that we have

$$\sup_{1\leq\eta\leq d/k} 2\sqrt{\eta(1 - k\eta/d)} + \eta/d < 1,$$

Since this holds providing that $k \geq (3d + 2)/4$, the proof of Theorem 16 is complete.

Lecture 6—Sums of Powers, and Parameterizations

In this lecture we shall look at the projective surface

$$x_1^d + x_2^d = x_3^d + x_4^d \tag{6.1}$$

and the affine surface

$$x_1^d + x_2^d + x_3^d = N.$$

It transpires that our basic techniques are well adapted to these, but need to be supplemented by information about possible curves of low degree (or low genus) lying on these varieties.

6.1 Theorem 13—Equal Sums of Two Powers

We shall begin by examining the surface (6.1). We can make a direct application of Theorem 14 to show that all points in the box $\max |x_i| \leq B$ must lie on one of $O_{\varepsilon,d}(B^{3\sqrt{d}+\varepsilon})$ curves in the surface. Moreover Theorem 12 tells us that $O_d(1)$ of these curves can have degree $d - 2$ or lower.

We begin by disposing of points which lie on a curve C of degree $k \geq d-1$ in the surface (6.1). Here we can apply Theorem 4 to show that any such curve contributes $O_{\varepsilon,k}(B^{2/k+\varepsilon})$ to $N_1(F;B)$. Since we have $d-1 \leq k \ll_{d,\varepsilon} 1$ it follows that the total contribution to $N_1(F;B)$ from all such curves arising from Theorem 14 will be $O_\varepsilon(B^{3/\sqrt{d}+2/(d-1)+\varepsilon})$, as required for Theorem 13.

We therefore turn our attention to the curves of degree at most $d-2$. Theorem 12 assures us that we can produce a finite list of these, independently of B. Thus we do not have the usual problem of uniformity with respect to B. It follows that we may apply the theorems (1.3) and (1.4) of Néron and Faltings to show that any curve of genus 1 will contribute $O_{\varepsilon,d}(B^\varepsilon)$ and any curve of genus 2 or more will contribute $O_d(1)$. Thus we have a total contribution of $O_{\varepsilon,d}(B^\varepsilon)$ to $N_1(F;B)$.

It follows that we must now examine the possibility that there are curves of genus zero on the surface (6.1). In doing this it in fact suffices to work over $\mathbb{C}$. Since any curve of genus zero can then be parameterized by polynomials we shall look for possible polynomial solutions to (6.1). This is clearly a question of interest in its own right, in view of the solutions (1.6) and (1.8) for the cases $d=3$ and $d=4$.

We shall establish the following general result, following an argument due to Newman and Slater [26].

Lemma 5 *Let $n \geq 2$ and let $f_1(t), \dots, f_n(t) \in \mathbb{C}[t]$ be non-zero polynomials. Suppose that $d \geq n(n-2)$. Then if*

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

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

If one of the polynomials is constant it suffices to have $d \geq (n-2)(n-1)$.

We therefore see that there can be no analogue of Euler's parametric solution to (1.7) for the surfaces (6.1) when the degree is 8 or more. Indeed, when $n=4$ and $d \geq 8$ we may conclude that $f_4(t) = cf_3(t)$, say. Then either $c^d = -1$, or

$$f_1(t)^d + f_2(t)^d + \tilde{f}_3(t)^d = 0$$

with

$$\tilde{f}_3(t) = (1+c^d)^{1/d} f_3(t) \neq 0.$$

In the first case we must have $f_1(t)^d + f_2(t)^d = f_3(t)^d + f_4(t)^d = 0$. In the second case, Lemma 5 shows that at least two of $f_1(t), f_2(t), \tilde{f}_3(t)$ are proportional, and hence that all three are. In either case we see that the original polynomials $f_1(t), \dots, f_4(t)$ are all proportional. These polynomials would then not parameterize a curve. We therefore see that, for $d \geq 8$, any curve of genus zero lying in the surface

$$x_1^d + x_2^d - x_3^d - x_4^d = 0$$

must be one of the obvious lines, and hence cannot contribute to $N_1(F;B)$. This establishes Theorem 13 when $d \geq 8$. On the other hand, if $d \leq 7$ one has $3/\sqrt{d} + 2/(d-1) \geq 1$, so that Theorem 13 follows from Theorem 11 in this case.

We now prove Lemma 5, which will be done by induction on n, the result being trivial for $n = 2$. Clearly we can suppose that the polynomials $f_j(t)$ have no common factor. It will be convenient to write $F_j(t) = f_j(t)^d$. We begin by differentiating the relation

$$\sum_{j=1}^{n} F_j(t) = 0$$

repeatedly, and we set

$$H_{ij}(t) = F_j^{-1}(t)\big(\frac{d}{dt}\big)^i F_j(t) \quad (0 \le i \le n-2). \tag{6.2}$$

We then deduce a system of equations

$$\sum_{j=1}^{n} H_{ij}(t)F_j(t) = 0 \quad (0 \le i \le n-2),$$

which we write in matrix form as $H\mathbf{F} = \mathbf{0}$, where H is the $(n-1) \times n$ matrix with entries $H_{ij}(t)$, and $\mathbf{F}$ is the column vector of length n, with entries $F_j(t)$.

We consider two cases. Suppose firstly that H has rank strictly less than $n-1$. In this case all the $(n-1) \times (n-1)$ minors $H_j(t)$ (say) must vanish. We now observe that

$$F_1(t)F_2(t)\dots F_{n-1}(t)H_n(t) = \begin{vmatrix} F_1(t) & \dots & F_{n-1}(t) \\ \dots & & \dots \\ \big(\frac{d}{dt}\big)^{n-2}F_1(t) & \dots & \big(\frac{d}{dt}\big)^{n-2}F_{n-1}(t) \end{vmatrix},$$

which is the Wronskian of $F_1(t), \dots, F_{n-1}(t)$. According to our assumption this vanishes, and hence the polynomials $F_1(t), \dots, F_{n-1}(t)$ will be linearly dependent over $\mathbb{C}$. There is therefore a set $\mathcal{S} \subseteq \{1, \dots, n-1\}$ for which we have a relation

$$\sum_{j \in \mathcal{S}} \alpha_j F_j(t) = 0$$

in which none of the α_j are zero. Moreover, if $\#\mathcal{S} = n'$, we will have $2 \le n' \le n-1$. We can now pick any d-th roots $\alpha^{1/d}$ to obtain an equation of the form

$$\sum_{j \in \mathcal{S}} \{\alpha_j^{1/d} f_j(t)\}^d = 0.$$

According to our induction hypothesis, two of the polynomials f_i, f_j must be proportional, as required.

We turn now to the second case, in which the rank of H is $n-1$. In this case $\mathbf{F}$ must be proportional to $(H_1, \dots, H_n)$. Without loss of generality we can assume that the degree h, say, of $f_1(t)$ is maximal. On recalling that $F_j(t) = f_j(t)^d$ we see from (6.2) that there are polynomials $g_{ij}(t)$ such that

$$H_{ij}(t) = \frac{g_{ij}(t)}{f_j(t)^i}, \quad \deg(g_{ij}(t)) \le i(h-1).$$

consequently, if we define

$$Q(t) = \{\prod_{j=1}^{n} f_j(t)\}^{n-2},$$

and set

$$Q(t)H_j(t) = P_j(t), \quad (1 \leq j \leq n), \tag{6.3}$$

it follows that $P_j(t)$ must be a polynomial, and that

$$\deg(P_j(t)) \leq (n-2)nh - \frac{(n-1)(n-2)}{2}. \tag{6.4}$$

However the polynomials $F_1(t), \ldots, F_n(t)$ will be coprime, and $\mathbf{F}(t)$ is proportional to $(P_1(t), \ldots, P_n(t))$. It follows that

$$hd = \deg(f_1(t)^d) = \deg(F_1(t)) \leq \deg(P_1(t)) \leq (n-2)nh - \frac{(n-1)(n-2)}{2}.$$

We therefore obtain a contradiction if

$$d > (n-2)(n - \frac{n-1}{2h}).$$

In particular we cannot have $d \geq n(n-2)$, irrespective of the value of h. If one of the polynomials is constant then the bound (6.4) may be replaced by

$$\deg(P_j(t)) \leq (n-2)(n-1)h - \frac{(n-1)(n-2)}{2},$$

and we cobtain a contradiction when $d \geq (n-2)(n-1)$. This completes the proof of Lemma 5.

6.2 Parameterization by Elliptic Functions

Lemma 5 gives good control over possible genus zero curves on diagonal hypersurfaces

$$X_1^d + \ldots + X_n^d = 0. \tag{6.5}$$

One can prove analogous results for genus 1 curves in general, but here we shall restrict attention to plane cubic curves. These can be parameterized using the Weierstrass elliptic function. Specifically, if there is a plane cubic curve contained in the variety (6.5), then there are functions

$$f_j(z) = A_j \wp'(z) + B_j \wp(z) + C_j, \quad (1 \leq j \leq n), \tag{6.6}$$

not all proportional to each other, such that

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

identically for $z \in \mathbb{C}$. Since the vector $(f_1(z), \ldots, f_n(z))$ must describe a cubic curve rather than a straight line, we conclude that the matrix

$$M = \begin{pmatrix} A_1 & \ldots & A_n \\ B_1 & \ldots & B_n \\ C_1 & \ldots & C_n \end{pmatrix}$$

must have rank 3,

We shall now have the following result, analogous to Lemma 5.

Lemma 6 *Let $n \geq 2$ and let $f_1(z), \ldots, f_n(z)$ satisfy (6.6). Assume further that the corresponding matrix M has rank 3, and that $d > (7n-1)(n-2)/6$. 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.

Before proving this we note that Green [9] gives a result which is both stronger and more general than Lemma 6. He proves that it suffices to have $d \geq (n-1)^2$, even when the f_j are arbitrary meromorphic functions which do not all vanish simultaneously. (We will also encounter such a condition, but it causes no problem for functions of the form (6.6).) Green's argument uses Nevanlinna theory, while our proof is more explicit and self-contained.

The result given in Lemma 6 is trivial if any function $f_j(z)$ vanishes identically, so we shall assume that each such function is non-zero. At least one matrix

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

where $i < j < k$, will be non-singular, since M has rank 3. It follows that the simultaneous equations

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

have no solution. Hence there can be no $z_0 \in \mathbb{C}$ at which $f_i(z), f_j(z), f_k(z)$ all vanish.

The argument now follows that given for Lemma 5. If the matrix H has rank $n-2$ or less, there will be two functions which are proportional. On the other hand, if the rank of H is $n-1$, then the vector $\mathbf{F}$ will be proportional to $(H_1, \ldots, H_n)$. In particular we see that none of the H_j can vanish identically, so that we may write

$$\frac{F_i(z)}{F_j(z)} = \left(\frac{f_i(z)}{f_j(z)}\right)^d = \frac{H_i(z)}{H_j(z)} = \frac{P_i(z)}{P_j(z)}, \tag{6.7}$$

with functions $P_i(z), P_j(z)$ constructed as in (6.3). This construction shows that these functions are polynomials in $\wp$ and its derivatives, and that they have poles of order at most

$$3n(n-2) + \frac{(n-1)(n-2)}{2} = \frac{(7n-1)(n-2)}{2} \tag{6.8}$$

at the origin.

In completing the proof we shall use the fact that a doubly periodic meromorphic function has the same (finite) number of poles as zeros, counted according to multiplicity, in its fundamental parallelogram. There must be some index i for which $A_i \neq 0$, since M has rank 3. For this index, $f_i(z)$ has a single pole, of

order 3. Thus f_i has zeros of total multiplicity 3. Suppose that z_0 is such a zero, and has multiplicity μ say. We have already noted that the functions $f_1, \dots, f_n$ cannot all vanish at z_0. There is therefore an index j with $f_j(z_0) \neq 0$. Since P_j is a polynomial in $\wp$ and its derivatives, it has poles only at the origin, within the fundamental parallelogram. We then see from (6.7) that P_i will have a zero of order at least μd at the point z_0. We may apply this reasoning to each of the zeros of f_i and show that P_i has zeros of total multiplicity at least $3d$ in the fundamental parallelogram. On the other hand, (6.8) provides an upper bound for the multiplicity of the poles of P_i. A comparison of these bounds yields

$$3d \leq 3n(n-2) + \frac{(n-1)(n-2)}{2},$$

contradicting the assumption of the lemma. This suffices for the proof.

6.3 Sums of 3 Powers

We turn now to the affine surface

$$x_1^d + x_2^d + x_3^d = N. \tag{6.9}$$

Bearing in mind the arithmetical significance of this we will only look at solutions with $x_i > 0$. Let $r(N)$ be the number of such solutions. When $d \geq 2$ we easily have $r(N) \ll_{d,\varepsilon} N^{1/d+\varepsilon}$, but no improvement in the exponent $1/d$ has hitherto been given, for any value of d. The exponent is certainly best possible for $d = 2$, and for $d = 3$ it was shown by Mahler [24] that $r(N) = \Omega(N^{1/12})$. This follows by taking $N = n^{12}$ in the identity

$$(9x^4)^3 + (3xn^3 - 9x^4)^3 + (n^4 - 9x^3n)^3 = n^{12}. \tag{6.10}$$

In general such an identity must arise from an expression of a non-zero constant as a sum of three d-th powers of polynomials. A comparison of the leading terms in such an identity shows that d must be odd, while Lemma 5 shows that we must have $d < 6$. Thus there can be no analogue of Mahler's identity for higher powers, except possibly for $d = 5$. Indeed it may be conjectured that $r(N) \ll_{d,\varepsilon} N^{\varepsilon}$ as soon as $d \geq 4$. The following result goes some way towards this.

Theorem 17 *For $d \geq 8$ we have*

$$r(N) \ll_{\varepsilon} N^{\theta/d+\varepsilon}$$

where

$$\theta = \frac{2}{\sqrt{d}} + \frac{2}{d-1}.$$

Observe that we have a non-trivial bound $\theta < 1$ for $d \geq 8$. Note also that the theorem remains true for $d < 8$, by the trivial bound $r(N) \ll_{d,\varepsilon} N^{1/d+\varepsilon}$, since $\theta > 1$ for $d < 8$.

For the proof we begin by applying Theorem 15 to the polynomial

$$F(x_1, x_2, x_3) = x_1^d + x_2^d + x_3^d - N,$$

taking $B_1 = B_2 = B_3 = N^{1/d} = B$, say. We conclude that all relevant points lie on one of $O_\varepsilon(B^{2/\sqrt{d}+\varepsilon})$ curves, each having degree $O_\varepsilon(1)$. When such a curve has degree $D \geq d-1$, it will have $O_\varepsilon(B^{2/D+\varepsilon})$ corresponding points, by Theorem 4. The total number of solutions of (6.9) in such cases is thus $O_\varepsilon(B^{\theta+\varepsilon})$, which is satisfactory. Indeed for curves in $\mathbb{A}^3$ Pila [28, Theorem A] shows that one may replace the exponent $2/d$ in our Theorem 4 by $1/d$. Thus in our situation each curve of degree $D \geq d-1$ will contribute $O_\varepsilon(B^{1/D+\varepsilon})$. It follows that Pila's result allows us to improve θ to

$$\theta = \frac{2}{\sqrt{d}} + \frac{1}{d-1}.$$

We have only stated the slightly weaker result in our theorem in order to be self-contained.

Now let C be a curve of degree at most $d-2$, contained in the surface (6.9). If $\theta : \mathbb{A}^3 \to \mathbb{A}^3$ is the map

$$\theta(x_1, x_2, x_3) = N^{-1/d}(x_1, x_2, x_3),$$

then $\theta(C)$ is a curve of degree at most $d-2$, lying in the non-singular surface S, given by $y_1^d + y_2^d + y_3^d = 1$. Theorem 12 has a natural affine version, which shows that there are $O(1)$ such curves, $C_1, \dots, C_t$, say. Obviously t and the curves C_i depend only on d and not on N. Suppose that $\theta(C) = C_i$, say. Let $\pi : \mathbb{A}^3 - \{\mathbf{0}\} \to \mathbb{P}^2$ be the map given by $\pi(x_1, x_2, x_3) = [(x_1, x_2, x_3)]$, where $[(x_1, x_2, x_3)]$ is the point in $\mathbb{P}^2$ represented by (x_1, x_2, x_3). Since $(0, 0, 0)$ does not satisfy (6.9) it cannot lie on the curve C, and hence π gives a regular map from C into $\mathbb{P}^2$. Moreover, since $\underline{\pi(\mathbf{x}) = \pi(\theta(\mathbf{x}))}$ we find that $\pi(C) = \pi(\theta(C)) = \underline{\pi(C_i)}$. Thus the Zariski closure $\overline{\pi(C)}$ must be one of a finite number of curves $\overline{\pi(C_i)}$, independent of N. Write $\Gamma_i = \overline{\pi(C_i)}$, for convenience.

This is the key point in our argument. The curves C are likely to be different for different values of N, and as N varies we will encounter infinitely many different curves C. Thus it appears that we will have a problem of uniformity in N. However these curves are all 'twists', by $N^{1/d}$, of a finite number of curves C_i, and by mapping into $\mathbb{P}^2$ each of these twists gets sent to the same curve Γ_i. The uniformity issue then disappears. Of course we are left with the problem that each point in $\mathbb{P}^2$ corresponds to many different points in $\mathbb{A}^3$. However, these are scalar multiples of one another, and at most one can be a solution of (6.9) for a particular value of N.

We now begin by disposing of the case in which Γ_i is not defined over $\mathbb{Q}$. In this case the rational point $\pi(\mathbf{x})$ lies on the intersection of Γ_i and any one of its conjugates. Such an intersection has $O(1)$ points by Bézout's Theorem. Thus Γ_i contains $O(1)$ rational points. As noted above, each such point in $\mathbb{P}^2$ can correspond to at most one solution of (6.9). Thus (6.9) has $O(1)$ solutions $\mathbf{x}$ for which $\pi(\mathbf{x})$ lies on a curve Γ_i not defined over $\mathbb{Q}$. Next, if Γ_i has genus 2 or more, it will have $O(1)$ points by Faltings' Theorem (1.4), and again there are $O(1)$ corresponding solutions of (6.9). In the case in which Γ_i has genus 1, Néron's result (1.4) similarly yields $O(B^\varepsilon)$ solutions of (6.9). Thus it remains to consider the case in which Γ_i is defined over $\mathbb{Q}$ and has genus zero.

In this final case we observe that a curve of genus zero defined over $\mathbb{Q}$ can be parameterized by rational functions. To be more precise, there are forms $f_1(u, v), f_2(u, v), f_3(u, v) \in \mathbb{Z}[u, v]$, with no common factor, such that every

rational point on Γ_i, with at most finitely many exceptions, is a non-zero rational multiple of $(f_1(u,v), f_2(u,v), f_3(u,v))$ for appropriate coprime $u, v \in \mathbb{Z}$. Thus it remains to examine solutions to the equation

$$\lambda^d\{f_1(u,v)^d + f_2(u,v)^d + f_3(u,v)^d\} = N, \quad \lambda \in \mathbb{Q}, \ \ u, v \in \mathbb{Z}, \ \ (u,v) = 1.$$

The forms f_1, f_2, f_3 can be considered fixed, independently of N, but λ, and of course u and v, may vary. We shall need to control λ. Write $\lambda = \mu/\nu$ with $(\mu, \nu) = 1$. We now use the fact that the forms f_i are coprime to produce relations of the form

$$\sum_{i=1}^{3} g_i(u,v) f_i(u,v) = Gu^r, \quad \sum_{i=1}^{3} h_i(u,v) f_i(u,v) = Hv^r,$$

where $g_i(u,v), h_i(u,v)$ are integral forms, and G, H are non-zero integer constants. Since $\lambda f_i(u,v)$ must be integral for $i = 1, 2, 3$, in any solution of interest, we conclude that $\nu | f_i(u,v)$, and hence that $\nu | Gu^r$ and $\nu | Hv^r$. However u and v are assumed to be coprime, so that $\nu | GH$. It follows that ν takes finitely many values. Moreover we have $\mu^d | N$, so that μ can take at most $O(N^\varepsilon)$ values.

We are left to consider the number of solutions of the Thue equation

$$f_1(u,v)^d + f_2(u,v)^d + f_3(u,v)^d = \nu^d \mu^{-d} N \tag{6.11}$$

for fixed forms f_1, f_2, f_3 and fixed μ, ν. We denote the form on the left by $F(u,v)$. If $F(u,v)$ has at least two distinct rational factors, say $F_1(u,v)$ and $F_2(u,v)$ then we will have $F_1(u,v) = N_1$ and $F_2(u,v) = N_2$ for certain factors N_1, N_2 of $\nu^d \mu^{-d} N$. These two equations determine $O(1)$ values of u, v, by elimination, so that (6.11) has $O_\varepsilon(N^\varepsilon)$ solutions. Now suppose to the contrary that F is a constant multiple of a power of an irreducible form F_1 say, in which case we have to consider solutions of an equation $F_1(u,v) = N_1$. When F_1 has degree 3 or more one may apply an old result of Lewis and Mahler [23], which shows that there are $O(A^{\omega(N_1)})$ such solutions, with a constant A depending only on F_1. This is enough to show that there are $O_\varepsilon(N^\varepsilon)$ solutions in this case.

Now consider the case in which F_1 has degree two. Here there will be $O_\varepsilon(N^\varepsilon)$ solutions to the equation $F_1(u,v) = N_1$ providing that we have $u, v \ll N$. However we assumed that the forms f_i had no common factor, and we may therefore take f_1, say, to be coprime to F_1. Now if $F_1(u,v) \ll N$ then there is a root α, say of $F_1(X,1) = 0$ such that $u - \alpha v \ll \sqrt{N} \ll N$. Similarly, since $f_1(u,v) \ll N$, we have $u - \beta v \ll N$ for some root β of $f_1(X,1)$. By subtraction we obtain $(\beta - \alpha)v \ll N$. Since F_1 and f_1 are coprime we will have $\alpha \neq \beta$, whence $v \ll N$. Similarly we have $u \ll N$. This provides the necessary bounds on u and v.

We have finally to consider the case in which F_1 is linear, so that $F(u,v) = c(au + bv)^k$, say. We then have an identity

$$f_1(u,v)^d + f_2(u,v)^d + f_3(u,v)^d = c(au + bv)^k,$$

and it is clear that d must divide k. But then Lemma 5 applies, since $d \geq 8$. Thus at least two of the terms must be proportional. A second application of the lemma then shows either that all four terms are proportional, contradicting the coprimality of f_1, f_2 and f_3, or that $f_i^d + f_j^d$ vanishes identically for some

pair of indices $i \neq j$. In the latter case the corresponding solutions to (6.9) cannot involve strictly positive integers. This completes the proof of Theorem 17.

Acknowledgements

These lecture notes were prepared while the author was visiting the Max-Planck Institute for Mathematics in Bonn. The hospitality and financial support of the Institute is gratefully acknowledged. Thanks are also due to Dr Browning and to Professors Perelli and Viola for their meticulous checking of these notes, the early versions of which contained numerous errors.

References

[1] M.A. Bennett, N.P. Dummigan and T.D. Wooley, The representation of integers by binary additive forms, *Compositio Math.,* 111 (1998), 15-33.

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

[3] T.D. Browning, Equal sums of two kth powers, *J. Number Theory*, 96 (2002), 293-318.

[4] N.D. Elkies, On $A^4 + B^4 + C^4 = D^4$, *Math. Comp.,* 51 (1988), 825-835.

[5] N.D. Elkies, Rational points near curves and small nozero $|x^3 - y^2|$ via lattice reduction, *Algorithmic number theory (Leiden, 2000), 33-63, Lecture Notes in Comput. Sci.,* 1838, (Springer, Berlin, 2000).

[6] P. Erdős, Arithmetical properties of polynomials, *J. London Math. Soc.,* 28, (1953). 416-425

[7] P. Erdős and K. Mahler, On the number of integers that can be represented by a binary form, *J. London Math. Soc.,* 13 (1938), 134-139.

[8] G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.,* 73 (1983), 349-366.

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

[10] J. Harris, *Algebraic geometry, A first course*, Graduate Texts in Mathematics, 133. (Springer-Verlag, New York, 1992).

[11] D.R. Heath-Brown, Cubic forms in ten variables, *Proc. London Math. Soc. (3),* 47 (1983), 225-257.

[12] D.R. Heath-Brown, The density of rational points on non-singular hypersurfaces, *Proc. Indian Acad. Sci. (Math. Sci.),* 104 (1994), 13-29.

[13] D.R. Heath-Brown, The density of rational points on cubic surfaces, *Acta Arithmetica,* 79 (1997), 17-30 .

[14] D.R. Heath-Brown, Counting rational points on cubic surfaces, *Astérisque,* 251 (1998), 13-29.

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

[16] C. Hooley, On binary cubic forms, *J. Reine Angew. Math.* 226 (1967), 30-87.

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

[18] C. Hooley, On the numbers that are representable as the sum of two cubes, *J. Reine Angew. Math.,* 314 (1980), 146-173.

[19] C. Hooley, On another sieve method and the numbers that are a sum of two h-th powers, *Proc. London Math. Soc. (3),* 43 (1981), 73-109.

[20] C. Hooley, On binary quartic forms, *J. Reine Angew. Math.,* 366 (1986), 32-52.

[21] C. Hooley, On another sieve method and the numbers that are a sum of two h-th powers. II, *J. Reine Angew. Math.,* 475 (1996), 55-75.

[22] C. Hooley, On binary cubic forms: II, *J. Reine Angew. Math.,* 521 (2000), 185-240.

[23] D.J. Lewis and K. Mahler, On the representation of integers by binary forms, *Acta Arith.,* 6 (1960/61), 333-363.

[24] K. Mahler, A note on hypothesis K of Hardy and Littlewood, *J. London Math. Soc.,* 11 (1936), 136-138.

[25] M. Nair, Power free values of polynomials, *Mathematika,* 23 (1976), 159-183.

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

[27] A.M. Ostrowski, Über die Bedeutung der Theorie der konvexen Polyheder für formale Algebra, *Jahresber. Deutsche Math.-Verein.,* 30 (1921), 98-99.

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

[29] W.M. Schmidt, Integer points on hypersurfaces, *Monatsh. Math.,* 102 (1986), 27-58.

[30] C. M. Skinner and T.D. Wooley, Sums of two k-th powers, *J. Reine Angew. Math.,* 462 (1995), 57-68.