13

Cubic forms

We now set out to prove that a homogeneous cubic equation

$$C(x_1, \ldots, x_n) = 0, \tag{13.1}$$

with integral coefficients, is always soluble in integers $x_1, \ldots, x_n$ (not all 0) if $n \geq 17$. The first such result, with the condition $n \geq 32$, was proved in 1957 [21], and the improved result was found early in 1962 [20]. In 1963 I proved that the condition $n \geq 16$ suffices [22], but this requires a more detailed argument of a somewhat special nature, beyond what is needed for 17.

It was pointed out by Mordell [61] in 1937 that there exist cubic forms in nine variables which do not represent zero, and consequently the condition $n \geq 10$ is essential if (13.1) is to be always soluble. The example of Mordell is based on the properties of a norm form of a cubic field. If p is a prime which does not factorize in the field, then the norm form $N(x, y, z)$ is never divisible by p except when x, y, z are all divisible by p. It follows easily that the equation

$$N(x_1, x_2, x_3) + pN(x_4, x_5, x_6) + p^2 N(x_7, x_8, x_9) = 0$$

has no solution in integers $x_1, \ldots, x_9$ except the trivial solution. Indeed, we could assert further that the corresponding congruence to the modulus p^3 has no solution except with all the variables 0 (mod p). A simple example would be provided by taking $p = 7$ and

$$N(x, y, z) = x^3 + 2y^3 + 4z^3 - 6xyz,$$

this being the norm form of the field generated by $\sqrt[3]{2}$. A similar construction to that above gives examples of homogeneous equations of degree k in k^2 variables which are insoluble.

The proof of the theorem on cubic equations falls into several chapters, each of which is largely self-contained. We begin by considering the exponential sum associated with a cubic form. Write

$$C(\mathbf{x}) = C(x_1, \ldots, x_n) = \sum_i \sum_j \sum_k c_{ijk} x_i x_j x_k,$$

where the summations go from 1 to n. The coefficients c_{ijk} are integers, and we can suppose that c_{ijk} is a symmetrical function of i, j, k. Let P be a large positive integer. Let $\mathfrak{B}$ be a fixed box in n dimensional space, namely the cartesian product of n intervals

$$x_j' < x_j \le x_j'', \quad (1 \le j \le n).$$

We shall suppose, merely for convenience, that $x_j'' - x_j' < 1$. Let

$$S(\alpha) = \sum_{P\mathfrak{B}} e\left(\alpha C(x_1, \ldots, x_n)\right),$$

where the summation is over all integer points in the box $P\mathfrak{B}$, given by

$$Px_j' < x_j \le Px_j'', \quad (1 \le j \le n).$$

Let $\mathcal{N}(P)$ denote the number of integer points $\mathbf{x}$ in $P\mathfrak{B}$ which satisfy (13.1). Then

$$\mathcal{N}(P) = \int_0^1 S(\alpha)\, d\alpha. \tag{13.2}$$

Our aim (in principle) is to prove that, with a suitable choice of the box $\mathfrak{B}$, there is an asymptotic formula for $\mathcal{N}(P)$ as $P \to \infty$, in which the main term is of order P^{n-3}. Actually this is not always true. What we *shall* arrive at is the apparently paradoxical result that the asymptotic formula holds on the hypothesis that (13.1) is insoluble (so that $\mathcal{N}(P) = 0$)! This will suffice for our purpose, because it will prove that (13.1) is soluble.

The trivial estimate for $|S(\alpha)|$ is P^n. We begin by investigating what happens if, for some particular α,

$$|S(\alpha)| \ge P^{n-K}, \tag{13.3}$$

where K is some positive number. Ultimately our aim, as in earlier chapters, is to be able to remove from the set of α any subset which contributes an amount of lower order than P^{n-3} to the integral.

The first step is to prove a generalization of Weyl's inequality. We define, for any two points $\mathbf{x}$, $\mathbf{y}$, a set of n bilinear forms:

$$B_j(\mathbf{x} \,|\, \mathbf{y}) = \sum_i \sum_k c_{ijk} x_i y_k, \quad (1 \le j \le n).$$

Lemma 13.1. *The hypothesis (13.3) implies that*

$$\sum_{|\mathbf{x}|<P}\sum_{|\mathbf{y}|<P}\prod_{j=1}^{n}\min\left(P, \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\|^{-1}\right) \gg P^{3n-4K}.$$

Proof. We have

$$\begin{aligned}|S(\alpha)|^2 &= \sum_{\mathbf{z}\in P\mathfrak{B}}\sum_{\mathbf{z}'\in P\mathfrak{B}} e\left(\alpha C(\mathbf{z}') - \alpha C(\mathbf{z})\right)\\ &= \sum_{\mathbf{z}\in P\mathfrak{B}}\sum_{\mathbf{y}+\mathbf{z}\in P\mathfrak{B}} e\left(\alpha C(\mathbf{y}+\mathbf{z}) - \alpha C(\mathbf{z})\right).\end{aligned}$$

For any $\mathbf{z}$, the box $P\mathfrak{B} - \mathbf{z}$ is contained in $|\mathbf{y}| < P$. Hence

$$|S(\alpha)|^2 \le \sum_{|\mathbf{y}|<P}\left|\sum_{\mathbf{z}\in\mathcal{R}(\mathbf{y})} e\left(\alpha C(\mathbf{y}+\mathbf{z}) - \alpha C(\mathbf{z})\right)\right|,$$

where $\mathcal{R}(\mathbf{y})$ denotes the common part of the boxes $P\mathfrak{B}$ and $P\mathfrak{B} - \mathbf{y}$. By Cauchy's inequality,

$$|S(\alpha)|^4 \ll P^n \sum_{|\mathbf{y}|<P}\left|\sum_{\mathbf{z}\in\mathcal{R}(y)} e\left(\alpha C(\mathbf{y}+\mathbf{z}) - \alpha C(\mathbf{z})\right)\right|^2.$$

We now repeat the argument on the inner sum. Its square does not exceed

$$\sum_{|\mathbf{x}|<P}\left|\sum_{\mathbf{z}\in\mathcal{S}(\mathbf{x},\mathbf{y})} e\left(\alpha C(\mathbf{z}+\mathbf{x}+\mathbf{y}) - \alpha C(\mathbf{z}+\mathbf{x}) - \alpha C(\mathbf{z}+\mathbf{y}) + \alpha C(\mathbf{z})\right)\right|,$$

where $\mathcal{S}(\mathbf{x},\mathbf{y})$ is a box for $\mathbf{z}$, depending on $\mathbf{x}$ and $\mathbf{y}$, with edges less than P in length. We have

$$\begin{aligned}&\alpha(C(\mathbf{z}+\mathbf{x}+\mathbf{y}) - C(\mathbf{z}+\mathbf{x}) - C(\mathbf{z}+\mathbf{y}) + C(\mathbf{z}))\\ &\quad= 6\alpha\sum_{i,j,k} c_{ijk}x_i y_k z_j + \phi = 6\alpha\sum_{j} z_j B_j(\mathbf{x}\,|\,\mathbf{y}) + \phi,\end{aligned}$$

where ϕ does not involve z. By a now familiar estimate,

$$\left|\sum_{\mathbf{z}\in\mathcal{S}(\mathbf{x},\mathbf{y})} e\left(6\alpha\sum_j B_j(\mathbf{x}\,|\,\mathbf{y})z_j\right)\right| \ll \prod_{j=1}^{n}\min\left(P, \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\|^{-1}\right).$$

Substitution in the previous inequalities, together with (13.3), yields the result. $\square$

Note. It is a useful precaution, when estimating an exponential sum, to see what the trivial estimate yields, in order to judge how much (if anything) has been lost for ever. In the present case, taking P for the minimum throughout the product, the trivial estimate for $|S(\alpha)|^4$ would be P^{4n}, which is satisfactory.

Lemma 13.2. *The hypothesis (13.3) implies that the number of pairs* $\mathbf{x}$, $\mathbf{y}$ *of integer points that satisfy*[1]

$$|\mathbf{x}| < P, \quad |\mathbf{y}| < P, \quad \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\| < P^{-1}, \quad (1 \le j \le n), \tag{13.4}$$

is

$$\gg P^{2n-4K}(\log P)^{-n}.$$

Proof. Let $Y(\mathbf{x})$ denote the number of points $\mathbf{y}$ satisfying (13.4) for given $\mathbf{x}$. Then, for any integers $r_1, \ldots, r_n$ with $0 \le r_j < P$, there cannot be more than $Y(\mathbf{x})$ integer points $\mathbf{y}$, with each coordinate in some prescribed interval of length P, which satisfy

$$\frac{r_j}{P} \le \{6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\} < \frac{r_j+1}{P}, \quad (1 \le j \le n),$$

where $\{\theta\}$ denotes the fractional part of any real number θ. For if $\mathbf{y}'$ were one such point, and $\mathbf{y}$ were any such point, we should have $|\mathbf{y}-\mathbf{y}'| < P$ and

$$\|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y}-\mathbf{y}')\| < P^{-1}, \quad (1 \le j \le n).$$

Thus there cannot be more than $Y(\mathbf{x})$ possibilities for $\mathbf{y}$. (Note that $\mathbf{y} = \mathbf{0}$ is permitted in (13.4).)

Dividing the cube $|\mathbf{y}| < P$ into 2^n cubes of side P, we obtain

$$\begin{aligned}
&\sum_{|\mathbf{y}|<P} \prod_{j=1}^{n} \min(P, \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\|^{-1}) \\
&\ll Y(\mathbf{x}) \sum_{r_1=0}^{P-1} \cdots \sum_{r_n=0}^{P-1} \min\left(P, \frac{P}{r_j}, \frac{P}{r_j-1}\right) \\
&\ll Y(\mathbf{x})(P\log P)^n.
\end{aligned}$$

[1] Here we put $|\mathbf{x}| = \max(|x_1|, \ldots, |x_n|)$ for any point $\mathbf{x}$. This is a different notation from that of Chapter 12 where $|\mathbf{x}|$ denotes the distance of the point $\mathbf{x}$ from the origin. For our purposes, however, the difference is unimportant.

Substitution in the result of Lemma 13.1 gives

$$(P \log P)^n \sum_{|\mathbf{x}|<P} Y(\mathbf{x}) \gg P^{3n-4K}.$$

Since $\sum Y(\mathbf{x})$ is the number of pairs $\mathbf{x}$, $\mathbf{y}$ satisfying (13.4), the result follows. ☐

Note. Since the trivial estimate for the number of pairs $\mathbf{x}$, $\mathbf{y}$ satisfying (13.4) is P^{2n}, we have abandoned a factor $(\log P)^n$. But this is not important.

Lemma 13.3. *Let θ be independent of P and satisfy $0 < \theta < 1$. The hypothesis (13.3) implies that the number of pairs $\mathbf{x}$, $\mathbf{y}$ of integer points satisfying*

$$|\mathbf{x}| < P^\theta, \quad |\mathbf{y}| < P^\theta, \quad \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\| < P^{-3+2\theta} \tag{13.5}$$

is

$$\gg P^{2n\theta-4K}(\log P)^{-n}.$$

Proof. We shall get the result by two applications of Lemma 12.6. First we look upon $\mathbf{x}$ as fixed in (13.4), and consider the number of integer points $\mathbf{y}$. The inequalities for $\mathbf{y}$ are

$$\begin{cases} |y_1| < P, \dots, |y_n| < P, \\ |L_1(\mathbf{y}) - u_{n+1}| < P^{-1}, \dots, |L_n(\mathbf{y}) - u_{2n}| < P^{-1}, \end{cases} \tag{13.6}$$

where $L_j(\mathbf{y}) = 6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})$, and where u_{n+j} is the integer nearest to $L_j(\mathbf{y})$. The forms $L_j(\mathbf{y})$ satisfy the symmetry condition of the preceding chapter, since the coefficient of y_k in L_j is $6\alpha \sum_i c_{ijk} x_i$, and this is unaltered by interchanging j and k. We apply Lemma 12.6 with $y_1, \dots, y_n$ for $u_1, \dots, u_n$ and with

$$a = P, \quad Z_2 = 1, \quad Z_1 = P^{-1+\theta}.$$

When $Z_2 = 1$, the inequalities of Lemma 12.6 are the inequalities (13.6) above. Suppose these have $N(\mathbf{x})$ solutions in $\mathbf{y}$. The inequalities of Lemma 12.6 with $Z = Z_1$ become

$$|y_1| < P^\theta, \dots, |y_n| < P^\theta,$$
$$|L_1(\mathbf{y}) - u_{n+1}| < P^{-2+\theta}, \dots, |L_n(\mathbf{y}) - u_{2n}| < P^{-2+\theta}.$$

Hence the number of solutions of these in $\mathbf{y}$ is

$$\gg N(\mathbf{x}) P^{-n(1-\theta)}.$$

Hence the number of pairs $\mathbf{x}$, $\mathbf{y}$ which satisfy

$$|\mathbf{x}| < P, \quad |\mathbf{y}| < P^{\theta}, \quad \|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\| < P^{-2+\theta} \tag{13.7}$$

is

$$\gg P^{-n(1-\theta)} \sum_{|\mathbf{x}|<P} N(\mathbf{x}) \gg P^{n+n\theta-4K}(\log P)^{-n},$$

by Lemma 13.2.

Now we go through a similar argument, but with $\mathbf{y}$ fixed and $\mathbf{x}$ variable. For each $\mathbf{y}$, the conditions (13.7) on $\mathbf{x}$ are

$$|x_1| < P, \dots, |x_n| < P,$$
$$|M_1(\mathbf{x}) - u_{n+1}| < P^{-2+\theta}, \dots, |M_n(\mathbf{x}) - u_{2n}| < P^{-2+\theta},$$

where $M_j(\mathbf{x}) = B_j(\mathbf{y}\,|\,\mathbf{x})$ and u_{n+j} is now the integer nearest to $M_j(\mathbf{x})$. These are the inequalities of Lemma 12.6 with

$$a = P^{\frac{3}{2}-\frac{1}{2}\theta}, \quad Z = Z_2 = P^{-\frac{1}{2}+\frac{1}{2}\theta}.$$

We take $Z_1 = P^{-\frac{3}{2}+\frac{3}{2}\theta}$. Then the lemma tells us that the number of solutions of

$$|x_1| < P^{\theta}, \dots, |x_n| < P^{\theta},$$
$$|M_1(\mathbf{x}) - u_{n+1}| < P^{-3+2\theta}, \dots, |M_n(\mathbf{x}) - u_{2n}| < P^{-3+2\theta},$$

is $\gg P^{-n(1-\theta)}N_1(\mathbf{y})$, where $N_1(\mathbf{y})$ denotes the number of solutions of (13.7) in $\mathbf{x}$ for given $\mathbf{y}$. Hence the number of pairs of integer points satisfying (13.5) is

$$\gg P^{-n+n\theta} \sum_{|\mathbf{y}|<P^{\theta}} N_1(\mathbf{y}) \gg P^{-n+n\theta}P^{n+n\theta-4K}(\log P)^{-n},$$

whence the result. □

Lemma 13.4. *Let θ be independent of P and satisfy $0 < \theta < 1$. Let ε be any small fixed positive number. Then either*

(A) *there are more than $P^{n\theta+\varepsilon}$ pairs $\mathbf{x}$, $\mathbf{y}$ of integer points satisfying*

$$|\mathbf{x}| < P^{\theta}, \quad |\mathbf{y}| < P^{\theta}, \quad B_j(\mathbf{x}\,|\,\mathbf{y}) = 0, \quad (1 \le j \le n), \tag{13.8}$$

or

(B) *for every α the hypothesis*

$$|S(\alpha)| \ge P^{n-\frac{1}{4}n\theta+\varepsilon} \tag{13.9}$$

implies that α has a rational approximation a/q such that

$$(a,q)=1, \quad 1 \le q \ll P^{2\theta}, \quad |q\alpha - a| < P^{-3+2\theta}. \tag{13.10}$$

Proof. We take $K = \frac{1}{4}n\theta - \varepsilon$ in (13.3), so that this becomes the same as (13.9). By Lemma 13.3, there are

$$\gg P^{2n\theta-4K}(\log P)^{-n} \gg P^{n\theta+\varepsilon}$$

pairs $\mathbf{x}$, $\mathbf{y}$ of integer points satisfying (13.5). If, for all points, $B_j(\mathbf{x}\,|\,\mathbf{y}) = 0$ for all j, we have alternative (A) active. If not, then for some pair $\mathbf{x}$, $\mathbf{y}$ in (13.5) and some j, we have $B_j(\mathbf{x}\,|\,\mathbf{y}) \neq 0$ and

$$\|6\alpha B_j(\mathbf{x}\,|\,\mathbf{y})\| < P^{-3+2\theta}.$$

Take $q = 6|B_j(\mathbf{x}\,|\,\mathbf{y})|$ and take a to be the nearest integer to αq. Then

$$|q\alpha - a| < P^{-3+2\theta}.$$

Also

$$\begin{aligned} q \ll |B_j(\mathbf{x}\,|\,\mathbf{y})| &\ll \sum_{i,k} |c_{ijk}||x_i||y_k| \\ &\ll |\mathbf{x}||\mathbf{y}| \ll P^{2\theta}. \end{aligned}$$

We do not necessarily have $(a,q)=1$, but this can be ensured by removing any common factor from q and a. Thus we obtain alternative (B). $\square$

Note. It will be seen that alternative A does not involve α. Nor does it involve θ essentially, for if we put $R = P^\theta$, the assertion is that the number of pairs $\mathbf{x}$, $\mathbf{y}$ satisfying

$$|\mathbf{x}| < R, \quad |\mathbf{y}| < R, \quad B_j(\mathbf{x}\,|\,\mathbf{y}) = 0, \quad (1 \le j \le n), \tag{13.11}$$

is greater than $R^{n+\varepsilon'}$ for some fixed $\varepsilon' > 0$. Thus alternative A relates to an intrinsic property of the cubic form. Note that we could exclude $\mathbf{x} = \mathbf{0}$ and $\mathbf{y} = \mathbf{0}$ from (13.11) if we wished, since the number of such pairs is $\ll R^n$.

Alternative B gives us a situation which is similar in principle to that with which we have become familiar in earlier chapters. It will enable us, in Chapter 15, to estimate satisfactorily the contribution made to $\int_0^1 S(\alpha)\ d\alpha$ by a large set of α and leave us with a relatively small number of short intervals in which we can approximate to $S(\alpha)$.