14

Cubic forms: bilinear equations

We now investigate alternative A of Lemma 13.4, namely that for some arbitrarily large R there exist more than $R^{n+\varepsilon}$ pairs $\mathbf{x}, \mathbf{y}$ of integer points satisfying $0 < |\mathbf{x}| < R$, $0 < |\mathbf{y}| < R$ and

$$B_j(\mathbf{x} \,|\, \mathbf{y}) = 0, \quad (1 \leq j \leq n). \tag{14.1}$$

We have excluded $\mathbf{x} = \mathbf{0}$ and $\mathbf{y} = \mathbf{0}$, in accordance with the remark just made. In this chapter, which is self-contained, we shall prove that this implies the existence of an integer point $\mathbf{z} \neq \mathbf{0}$ for which $C(\mathbf{z}) = 0$. Actually a slightly weaker hypothesis would suffice, namely that there are more than AR^n pairs $\mathbf{x}, \mathbf{y}$, where A is greater than some function of n. But the use of ε simplifies the exposition.

For any particular $\mathbf{x}$, the equations (14.1) are n linear equations in $\mathbf{y} = (y_1, \ldots, y_n)$, and their determinant is

$$H(\mathbf{x}) = \det\left(\sum_{i=1}^{n} c_{ijk} x_i\right), \quad (1 \leq j, k \leq n). \tag{14.2}$$

This is the Hessian of the cubic form $C(\mathbf{x})$. It is a form of degree n in $\mathbf{x}$; or one should rather say of apparent degree n, since it may identically vanish. We must first prove that this cannot happen if $C(\mathbf{x})$ does not represent zero.

Lemma 14.1. *If $C(\mathbf{x}) \neq 0$ for all integral $\mathbf{x} \neq \mathbf{0}$, then $H(\mathbf{x})$ does not vanish identically.*

Proof. Suppose $H(\mathbf{x}) = 0$ identically. Let $n - r$ $(r \geq 1)$ be the identical rank of $H(\mathbf{x})$; that is, suppose all subdeterminants of the matrix in (14.2) of order $n - r + 1$ vanish identically in $\mathbf{x}$ but some subdeterminant of order $n - r$ does not. Suppose, for convenience of exposition, that the last determinant, say Δ, is in the top left-hand corner. Then the first $n-r$ of the equations (14.1) imply all the others, since the later rows of the matrix are linearly dependent on the first $n-r$ rows. Let Δ_j denote the determinant obtained from Δ by replacing the jth column by the $(n-r+1)$th column. Then a particular solution of the equations $B_j(\mathbf{x}\,|\,\mathbf{y}) = 0$ is given by

$$y_1 = \Delta_1, \ldots, y_{n-r} = \Delta_{n-r}, y_{n-r+1} = -\Delta, y_{n-r+2} = \cdots = 0.$$

(This follows from Cramer's rule: we solve the first $n-r$ equations, with $y_{n-r+1} = 1$, $y_{n-r+2} = 0, \ldots$, and multiply throughout by $-\Delta$.)

In this solution, $y_1, \ldots, y_n$ are forms in $x_1, \ldots, x_n$ with integral coefficients, which do not all vanish identically, since Δ does not. We have

$$B_j(\mathbf{x}\,|\,\mathbf{y}) = \sum_{i,k} c_{ijk} x_i y_k = 0$$

identically in $x_1, \ldots, x_n$. We now regard $x_1, \ldots, x_n$ as continuous variables, and differentiate this identity with respect to any x_ν, getting

$$\sum_k c_{\nu jk} y_k + \sum_{i,k} c_{ijk} x_i \frac{\partial y_k}{\partial x_\nu} = 0$$

for all j and ν. Multiply by y_j and sum over j, and note that

$$\sum_i \sum_j c_{ijk} x_i y_j = 0$$

for all k. We get

$$\sum_{j,k} c_{\nu jk} y_j y_k = 0$$

for all ν. This implies, in particular, that $C(\mathbf{y}) = 0$. If we now take $\mathbf{x}$ to be any integer point for which $\mathbf{y} \neq \mathbf{0}$ (as is possible because $\mathbf{y}$ is not identically $\mathbf{0}$) we get a contradiction to the hypothesis. This proves the lemma. $\square$

The last lemma shows that the points $\mathbf{x}$, for which there is a non-zero solution of the linear equations $B_j(\mathbf{x}\,|\,\mathbf{y}) = 0$ in $\mathbf{y}$, satisfy the non-identical equation $H(\mathbf{x}) = 0$. Thus the number of such integral points with $0 < |\mathbf{x}| < R$ is $\ll R^{n-1}$. We next extend this result by proving that, for $r = 1, \ldots, n-1$, the number of integral points $\mathbf{x}$ with $0 < |\mathbf{x}| < R$ for which the equations have r linearly independent solutions in $\mathbf{y}$ is $\ll R^{n-r}$. The result already proved is the case $r = 1$.

It is convenient to deal first with a question of elementary algebraic geometry.

Lemma 14.2. *Let*

$$f_1(\mathbf{x}), \ldots, f_N(\mathbf{x})$$

be forms with integral coefficients in $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *and suppose that* N *and the (common) degree of* $f_1, \ldots, f_N$ *are bounded in terms of* n. *Suppose that for some arbitrarily large* R, *there is a set* $\mathcal{X}$ *of integer points* $\mathbf{x}$ *satisfying*

$$0 < |\mathbf{x}| < R, \qquad f_1(\mathbf{x}) = 0, \ldots, f_N(\mathbf{x}) = 0,$$

and suppose that there are more than $R^{n-r+\varepsilon}$ *points in* $\mathcal{X}$. *Then, for some one of the points* $\mathbf{x}$ *of* $\mathcal{X}$, *there exist numbers* $T_{i\rho}$, $D_{\rho\nu}$, *such that*

$$\frac{\partial f_i}{\partial x_\nu} = \sum_{\rho=1}^{r-1} T_{i\rho} D_{\rho\nu},$$

for $i = 1, \ldots, N$ *and* $\nu = 1, \ldots, n$, *at that point.*

Note. The equations for $\partial f_i / \partial x_\nu$ are equivalent to the assertion that the rank of the matrix $\partial f_i / \partial x_\nu$ for $i = 1, \ldots, N$ and $\nu = 1, \ldots, n$, is $\leq r - 1$. This follows from elementary matrix theory.

Proof. The equations $f_1(\mathbf{x}) = 0, \ldots, f_N(\mathbf{x}) = 0$ define an algebraic variety in n-dimensional space (or $(n-1)$-dimensional projective space). Such a variety is expressible as a union of absolutely irreducible varieties, and the number of these, under the present hypotheses, is bounded in terms of n. Hence there is one of these absolutely irreducible varieties, say $\mathcal{V}$, which contains more than $R^{n-r+\frac{1}{2}\varepsilon}$ points of $\mathcal{X}$.

Associated with the absolutely irreducible variety $\mathcal{V}$, considered as a set of points in complex space, is its *dimension* s. We need only the following property of s: the irreducible variety $\mathcal{V}$ can be decomposed into a bounded number of parts, such that, on each part, s of $x_1, \ldots, x_n$ are independent variables and the other $n - s$ are single valued differentiable functions of them.

It follows that $\mathcal{V}$ contains $\ll R^s$ integer points $\mathbf{x}$ satisfying $|\mathbf{x}| < R$, since there are $\ll R^s$ possibilities for each of the s coordinates. Comparison with the earlier statement about the number of points of $\mathcal{X}$ on $\mathcal{V}$ shows that

$$s \geq n - r + 1.$$

Now consider the neighbourhood of any one point of $\mathcal{X}$ on $\mathcal{V}$. We can suppose that here $x_{s+1}, \dots, x_n$ are single valued and differentiable functions of $x_1, \dots, x_s$. Let $f(x_1, \dots, x_n)$ be any differentiable function which vanishes everywhere on $\mathcal{V}$. Then, differentiating the identity $f = 0$ with respect to x_ν, for $\nu = 1, \dots, s$, we get

$$f^{(\nu)} + f^{(s+1)} \frac{\partial x_{s+1}}{\partial x_\nu} + \cdots + f^{(n)} \frac{\partial x_n}{\partial x_\nu} = 0,$$

where $f^{(j)} = \partial f / \partial x_j$ when $x_1, \dots, x_n$ are independent variables. Thus for $\nu = 1, \dots, s$ we have

$$\begin{aligned} f^{(\nu)} &= \sum_{\rho=1}^{n-s} f^{(s+\rho)} \left(-\frac{\partial x_{s+\rho}}{\partial x_\nu} \right) \\ &= \sum_{\rho=1}^{n-s} f^{(s+\rho)} D_{\rho,\nu}, \end{aligned}$$

say. If we define $D_{\rho,\nu}$ for $\nu > s$ by

$$D_{\rho,\nu} = \begin{cases} 1 & \text{if } s + \rho = \nu, \\ 0 & \text{otherwise,} \end{cases}$$

the same relations holds for $\nu = s + 1, \dots, n$. Hence

$$f^{(\nu)}(\mathbf{x}) = \sum_{\rho=1}^{n-s} T_\rho(f) D_{\rho,\nu}$$

for $\nu = 1, \dots, n$, where $T_\rho(f) = f^{(s+\rho)}$. Note that the numbers $D_{\rho,\nu}$ are independent of f, and the numbers $T_\rho(f)$ are independent of ν. This proves the relations in the enunciation, since $n - s \le r - 1$, and they hold at any point on $\mathcal{V}$, and in particular at any of the points of $\mathcal{X}$ on $\mathcal{V}$. □

Lemma 14.3. *Suppose $C(\mathbf{x}) \neq 0$ for all integral $\mathbf{x} \neq \mathbf{0}$. Then the number of integer points $\mathbf{x}$ with $0 < |\mathbf{x}| < R$ for which the bilinear equations $B_j(\mathbf{x} \,|\, \mathbf{y}) = 0$ have exactly r linearly independent solutions in $\mathbf{y}$ is less than $R^{n-r+\varepsilon}$.*

Proof. The points in question are those for which the rank of the matrix $(\sum_i c_{ijk} x_i)$ is exactly r. It will suffice to consider the set $\mathcal{X}$ of integer points for which some particular subdeterminant of order $n - r$ is $\neq 0$ and all subdeterminants of order $n - r + 1$ are 0, and to prove that the number of points in $\mathcal{X}$ is less than $R^{n-r+\varepsilon}$. Suppose the number of points in $\mathcal{X}$ is $\geq R^{n-r+\varepsilon}$. Then we have the situation of Lemma 14.2, where $f_1(\mathbf{x}), \ldots, f_N(\mathbf{x})$ are all the subdeterminants of order $n-r+1$.

For any points $\mathbf{x}$ of $\mathcal{X}$ we can construct r linearly independent solutions $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(r)}$ of the bilinear equations, as in the proof of Lemma 14.1, by taking the coordinates of these points to be certain subdeterminants of order $n-r$. (In that proof, we needed only one solution, which we got by taking $y_{n-r+1} = -\Delta$, $y_{n-r+2} = \cdots = 0$; but the extension to r solutions is immediate.)

Now consider the values of $B_j(\mathbf{x} \,|\, \mathbf{y}^{(p)})$ when $\mathbf{x}$ is an arbitrary point (real or complex) and $\mathbf{y}^{(p)}$ is as above. This value will be a certain subdeterminant of order $n-r+1$ of the matrix mentioned at the beginning. (Sometimes it will be a subdeterminant of order $n-r+1$ with two identical rows, but that is immaterial.) Hence for *any* point $\mathbf{x}$ we have the identities

$$\sum_{i,k} c_{ijk} x_i y_k^{(p)} = \Delta_{j,p}(\mathbf{x}),$$

where $\Delta_{j,p}(\mathbf{x})$ is some subdeterminant of order $n-r+1$. Of course, all the $\Delta_{j,p}$ vanish if $\mathbf{x}$ is in $\mathcal{X}$.

In the above identities, $x_1, \ldots, x_n$ are independent variables. Differentiation with respect to x_ν gives

$$\sum_k c_{\nu jk} y_k^{(p)} + \sum_{i,k} c_{ijk} x_i \frac{\partial y_k^{(p)}}{\partial x_\nu} = \Delta_{j,p}^{(\nu)}(\mathbf{x}),$$

where the superscript (ν) on the right denotes a partial derivative. Multiply by $y_j^{(q)}$ $(1 \leq q \leq r)$ and sum over j. We get

$$\sum_{j,k} c_{\nu jk} y_k^{(p)} y_j^{(q)} + \sum_k \Delta_{k,q} \frac{\partial y_k^{(p)}}{\partial x_\nu} = \sum_j y_j^{(q)} \Delta_{j,p}^{(\nu)}(\mathbf{x}).$$

Now consider any point $\mathbf{Y}$ in the linear space of r dimensions generated by $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(r)}$, say

$$\mathbf{Y} = \sum_{p=1}^{r} K_p \mathbf{y}^{(p)}.$$

For this point, we have

$$\sum_{j,k} c_{\nu jk} Y_j Y_k + \sum_{p=1}^{r} \sum_{q=1}^{r} K_p K_q \sum_k \Delta_{k,q} \frac{\partial y_k^{(p)}}{\partial x_\nu}$$

$$= \sum_{p=1}^{r} \sum_{q=1}^{r} K_p K_q \sum_j y_j^{(q)} \Delta_{j,p}^{(\nu)}$$
$$= \sum_{p=1}^{r} \sum_j A_{j,p} \Delta_{j,p}^{(\nu)},$$

say. This holds for $\nu = 1, \ldots, n$.

Take the point $\mathbf{x}$ to be one of the points found in Lemma 14.2. For this point, which is fixed from now onwards, we have $\Delta_{k,q} = 0$ for all k and q, and

$$\Delta_{j,p}^{(\nu)} = \sum_{\rho=1}^{r-1} T_{j,p,\rho} D_{\rho,\nu}.$$

Hence

$$\sum_{j,k} c_{\nu jk} Y_j Y_k = \sum_{p=1}^{r} \sum_j A_{j,p} \sum_{\rho=1}^{r-1} T_{j,p,\rho} D_{\rho,\nu}.$$

Finally, multiply by $Y_\nu = \sum_{\sigma=1}^{r} K_\sigma y_\nu^{(\sigma)}$ and sum over ν. We get

$$C(\mathbf{Y}) = \sum_{p=1}^{r} \sum_j A_{j,p} \sum_{\sigma=1}^{r} \sum_{\rho=1}^{r-1} K_\sigma y_\nu^{(\sigma)} T_{j,p,\rho} D_{\rho,\nu}.$$

Now choose the numbers $K_1, \ldots, K_r$ to satisfy

$$\sum_{\sigma=1}^{r} K_\sigma \sum_\nu y_\nu^{(\sigma)} D_{\rho,\nu} = 0, \quad (1 \le \rho \le r-1).$$

These are $r-1$ linear equations in r unknowns, and so have a solution with $K_1, \ldots, K_r$ not all 0. Also, since the numbers $D_{\rho,\nu}$ can be supposed rational, we can take $K_1, \ldots, K_r$ to be integers. Hence $\mathbf{Y}$ is an integer point not $\mathbf{0}$, because $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(r)}$ are linearly independent since the point $\mathbf{x}$ is in $\mathcal{X}$. We get $C(\mathbf{Y}) = 0$, contrary to hypothesis, and this proves the result. □

Note that in the above proof, the choice of $K_1, \ldots, K_r$ did not involve the numbers $A_{j,p}$. Had it done so, the reasoning would have been fallacious, since these themselves depend on $K_1, \ldots, K_r$.

Lemma 14.4. *Alternative A of Lemma 13.4 implies that $C(\mathbf{x})$ represents 0.*

Proof. If $C(\mathbf{x})$ does not represent zero then Lemma 14.1 and Lemma 14.3

imply that there are $\ll R^{n-r+\varepsilon}$ points $\mathbf{x}$ such that there are exactly r linearly independent solutions of the bilinear equations in $\mathbf{y}$. Hence there are $\ll R^{n+\varepsilon}$ pairs $\mathbf{x}, \mathbf{y}$ with $0 < |\mathbf{x}| < R$, $0 < |\mathbf{y}| < R$ which satisfy the bilinear equations. But this contradicts alternative A, since we can take the present ε to be (say) half of the ε in alternative A. (Actually the present proof shows that the number of pairs $\mathbf{x}, \mathbf{y}$ is $\ll R^n$, as remarked earlier.) $\square$