Chapter 5

ROTH'S LEMMA

1. Introduction

Roth bases the proof of his theorem on a general property of polynomials which is to be proved in this chapter. This property is roughly as follows.

Let

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1}\dots\sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1}\dots x_m^{i_m} \not\equiv 0$$

be a polynomial in m variables, with integral coefficients which are not "too large" in absolute values. Assume that

$$\max\left(\frac{r_2}{r_1}, \frac{r_3}{r_2}, \dots, \frac{r_m}{r_{m-1}}\right)$$

is a "very small" positive number. Further let

$$\kappa_1 = \frac{P_1}{Q_1}, \dots, \kappa_m = \frac{P_m}{Q_m}$$

be m rational numbers written in their simplified forms for which both the maxima

$$H_1 = \max(|P_1|, |Q_1|), \dots, H_m = \max(|P_m|, |Q_m|)$$

and the quotients

$$\frac{\log H_2}{\log H_1}, \frac{\log H_3}{\log H_2}, \dots, \frac{\log H_m}{\log H_{m-1}}$$

are "very large". Then $A(x_1,\dots,x_m)$ cannot vanish to a "very high" order at $x_1=\kappa_1,\dots,x_m=\kappa_m$. (An exact formulation of Roth's Lemma will be given at the end of this chapter).

The main idea of the proof consists in an induction for m, the number of variables, the case $m=1$ being trivial. This induction uses a test for *linear independence of polynomials* in terms of the so-called *generalised Wronski determinants*.

2. Linear dependence and independence.

Let

$$f_\nu = f_\nu(x_1,\dots,x_m) \qquad (\nu = 1,2,\dots,n)$$

be n rational functions of m variables, with coefficients in a field K. The functions are said to be *linearly dependent* (or for short, *dependent*) over K if there are elements $c_1,\dots,c_n$ of K not all zero such that

$$c_1 f_1 + \ldots + c_n f_n \equiv 0$$

identically in $x_1, \ldots, x_m$. If no such elements exist, then the functions are called *linearly independent* (or for short, *independent*) over K. Evidently, *if* $f_1, \ldots, f_n$ *are independent, none of these functions can vanish identically.*

Assume, in particular, that the coefficients of $f_1, \ldots, f_n$ lie in the rational field Γ, and that these functions are dependent over the real field P. *Then the functions are also dependent over* Γ. For the identity $c_1 f_1 + \ldots + c_n f_n \equiv 0$ is equivalent to a finite system of linear equations

$$c_1 \phi_{1\sigma} + \ldots + c_n \phi_{n\sigma} = 0 \qquad (\sigma = 1, 2, \ldots, s)$$

for $c_1, \ldots, c_n$ with rational coefficients $\phi_{\nu\sigma}$. By the hypothesis the rank of the matrix of this system of equations is smaller than n. The system has therefore also a solution $c_1, \ldots, c_n$ in rational numbers not all zero, whence the assertion.

Conversely, if $f_1, \ldots, f_n$ *have rational coefficients and are independent over* Γ, *then they are also independent over* P.

3. Generalised Wronski determinants.

The letter D, with or without suffixes, will be used to denote differential operators of the form

$$\frac{\partial^{j_1 + \ldots + j_m}}{\partial x_1^{j_1} \ldots \partial x_m^{j_m}}$$

where $j_1, \ldots, j_m$ are non-negative integers. The sum $j_1 + \ldots + j_m$ of these integers is called the *order* of D. Thus the unit operator 1 has the order 0 because $j_1 = \ldots = j_m = 0$.

Let

$$f_\nu = f_\nu(x_1, \ldots, x_m) \qquad (\nu = 1, 2, \ldots, n)$$

be n rational functions with real coefficients, and let $D_1, \ldots, D_n$ be n differential operators such that

the order of D_ν *does not exceed* $\nu - 1$ $(\nu = 1, 2, \ldots, n)$.

The determinant

$$\begin{pmatrix} D_1 \ldots D_n \\ f_1 \ldots f_n \end{pmatrix} = \begin{vmatrix} D_1 f_1 & D_1 f_2 & \ldots & D_1 f_n \\ D_2 f_1 & D_2 f_2 & \ldots & D_2 f_n \\ \vdots & \vdots & & \vdots \\ D_n f_1 & D_n f_2 & \ldots & D_n f_n \end{vmatrix}$$

is called a *generalised Wronski determinant* or a *Wronskian.*

This Wronskian evidently vanishes identically when the operators $D_1, \ldots, D_n$ are not all distinct. It also vanishes identically if $f_1, \ldots, f_n$ are linearly dependent over the real field. For an identity

$$c_1 f_1 + \ldots + c_n f_n \equiv 0$$

implies the n identities

$$c_1 D_\nu f_1 + \dots + c_n D_\nu f_n \equiv 0 \qquad (\nu = 1,2,\dots, n).$$

If now $c_1,\dots, c_n$ are not all zero, then the determinant of this system of linear equations for $c_1,\dots, c_n$ vanishes, and this determinant is the Wronskian we are considering.

Let these two trivial cases be excluded. It will then be proved that at least one Wronskian of the given functions is not identically zero, at least when $f_1,\dots, f_n$ are polynomials.

4. The case of functions of one variable.

Let

$$f_\nu = f_\nu(x) \qquad (\nu = 1,2,\dots, n)$$

be n rational functions in one variable x which have real coefficients and are independent over the real field; thus, in particular,

$$f_n(x) \not\equiv 0.$$

There is only one Wronskian of these functions that does not vanish trivially, viz. that Wronskian which belongs to the operators

$$D_1 = 1,\ D_2 = \frac{d}{dx},\ D_3 = \frac{d^2}{dx^2},\dots,\ D_n = \frac{d^{n-1}}{dx^{n-1}}.$$

We show by induction for n that this Wronskian is in fact distinct from zero.

This is obvious for n=1 since then

$$\binom{D_1}{f_1} = f_1(x) \not\equiv 0.$$

Let therefore $n \geq 2$, and assume that the assertion has already been proved for n-1 functions.

Put

$$F_\nu(x) = \frac{d}{dx}\left(\frac{f_\nu(x)}{f_n(x)}\right) \qquad (\nu = 1,2,\dots, n-1).$$

These n-1 functions are still independent. For any equation

$$c_1 F_1 + \dots + c_{n-1} F_{n-1} = \frac{d}{dx}\left(\frac{c_1 f_1 + \dots + c_{n-1} f_{n-1}}{f_n}\right) \equiv 0$$

with real coefficients implies, on integrating, that

$$c_1 f_1 + \dots + c_{n-1} f_{n-1} \equiv -c_n f_n,$$

where c_n is a further real number, whence $c_1 = \dots = c_{n-1} = c_n = 0$ because $f_1,\dots, f_n$ are independent by hypothesis.

It follows then from the induction hypothesis that

$$\begin{pmatrix} D_1 \dots D_{n-1} \\ F_1 \dots F_{n-1} \end{pmatrix} \not\equiv 0 \quad \text{where} \quad D_\nu = \frac{d^{\nu-1}}{dx^{\nu-1}}.$$

Next one easily shows that, for any rational function g, identically

$$\begin{pmatrix} D_1 \dots D_n \\ f_1 g \dots f_n g \end{pmatrix} = g^n \begin{pmatrix} D_1 \dots D_n \\ f_1 \dots f_n \end{pmatrix}.$$

Here choose $g=f_n^{-1}$. Then in the Wronskian on the left-hand side all but the first element of the n-th column vanish, and this determinant reduces to

$$\begin{pmatrix} D_1 \dots D_n \\ f_1 g \dots f_n g \end{pmatrix} = \begin{pmatrix} D_1 \dots D_{n-1} \\ F_1 \dots f_{n-1} \end{pmatrix} \not\equiv 0.$$

Hence, finally,

$$\begin{pmatrix} D_1 \dots D_n \\ f_1 \dots f_n \end{pmatrix} = \begin{pmatrix} D_1 \dots D_{n-1} \\ F_1 \dots F_{n-1} \end{pmatrix} f_n^{-n} \not\equiv 0,$$

whence the assertion.

5. The general case.

From now on let

$$f_\nu(x_1, \dots, x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} f_{i_1 \dots i_m}^{(\nu)} x_1^{i_1} \dots x_m^{i_m} \qquad (\nu = 1, 2, \dots, n)$$

be n polynomials in $x_1, \dots, x_m$ that have real coefficients and are independent over the real field. We want to show that at least one of the Wronskians in these functions is not identically zero. In the special case $m=1$ this assertion has just been proved, even for the more general class of rational functions. To reduce the general case to this special one, denote by x a new variable, by g a positive integer exceeding all the degrees $r_1, \dots, r_m$, and put

$$\phi_\nu(x) = f_\nu(x, x^g, \dots, x^{g^{m-1}}) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} f_{i_1 \dots i_m}^{(\nu)} x^{i_1 + i_2 g + i_3 g^2 + \dots + i_m g^{m-1}}.$$

$$(\nu = 1, 2, \dots, n).$$

The exponents $i_1+i_2g+i_3g^2+\dots+i_mg^{m-1}$ of x may be considered as representations to the basis g, with $i_1, i_2, \dots, i_m$ as the digits; for by the choice of g these numbers may assume only the values $0, 1, 2, \dots, g-1$. Since there is only one representation of any integer to the basis g, it follows that no two terms in the multiple sum for $\phi_\nu(x)$ are constant multiples of the same power of x.

This implies that $\phi_1(x), \dots, \phi_n(x)$ likewise are independent over the real field. For let $c_1, \dots, c_m$ be real numbers such that $c_1\phi_1+\dots+c_n\phi_n \equiv 0$. By what has just been shown, this identity requires that

$$c_1 f_{i_1 \dots i_m}^{(1)} + \dots + c_n f_{i_1 \dots i_m}^{(n)} = 0 \qquad \text{for all suffixes } i_1, \dots, i_m.$$

But then $c_1 f_1 + \dots + c_n f_n \equiv 0$, whence $c_1 = \dots = c_n = 0$ by the assumed independence of $f_1, \dots, f_n$.

The result of §4 may then be applied to $\phi_1, \dots, \phi_n$, giving

$$W(x) = \begin{pmatrix} 1 & \frac{d}{dx} & \frac{d^2}{dx^2} & \dots & \frac{d^{n-1}}{dx^{n-1}} \\ \phi_1 & \phi_2 & \phi_3 & \dots & \phi_n \end{pmatrix} \not\equiv 0.$$

Denote now by δ the linear operator

$$\delta = \frac{\partial}{\partial x_1} + g x^{g-1}\frac{\partial}{\partial x_2} + g^2 x^{g^2-1}\frac{\partial}{\partial x_3} + \ldots + g^{m-1} x^{g^{m-1}-1}\frac{\partial}{\partial x_m},$$

and by the sign $(\ )^*$ the operation of substituting

$x, x^g, x^{g^2}, \ldots, x^{g^{m-1}}$ for $x_1, x_2, x_3, \ldots, x_m$, respectively.

In this notation, evidently

$$\phi_\nu(x) = \{f_\nu(x_1, \ldots, x_m)\}^* \quad \text{and} \quad \frac{d}{dx}\phi_\nu(x) = \{\delta f_\nu(x_1, \ldots, x_m)\}^*.$$

It follows that repeated differentiation of $\phi_\nu(x)$ leads to a relation

$$\frac{d^\mu}{dx^\mu}\phi_\nu(x) = \{(\psi_{\mu 1}(x) D_{\mu 1} + \psi_{\mu 2}(x) D_{\mu 2} + \ldots + \psi_{\mu N_\mu}(x) D_{\mu N_\mu}) f_\nu(x_1, \ldots, x_m)\}^*.$$

Here N_μ is a positive integer depending on μ; $\psi_{\mu 1}, \ldots, \psi_{\mu N_\mu}$ are polynomials in x; and $D_{\mu 1}, \ldots, d_{\mu N_\mu}$ are differential operators at most of order μ in the variables $x_1, \ldots, x_m$. The polynomials and the operators depend on μ, but not on the functions f_ν or ϕ_ν.

Replace now in the determinant $W(x)$ the terms $\frac{d^\mu}{dx^\mu}\phi_\nu(x)$ by their expressions in the derivatives of $f_\nu(x)$. Then each term in the determinant becomes a sum, and $W(x)$ takes the form

$$W(x) = \sum_{k_1=1}^{N_1} \cdots \sum_{k_{n-1}=1}^{N_{n-1}} \psi_{1k_1}(x) \ldots \psi_{n-1\,k_{n-1}}(x) \begin{pmatrix} 1 & D_{1k_1} & D_{2k_2} & \ldots & D_{n-1\,k_{n-1}} \\ f_1 & f_2 & f_3 & \ldots & f_n \end{pmatrix}^*.$$

Since $W(x) \not\equiv 0$, at least one of the terms on the right-hand side likewise is not identically zero, so that, say,

$$\begin{pmatrix} 1 & D_{1k_1} & D_{2k_2} & \ldots & D_{n-1\,k_{n-1}} \\ f_1 & f_2 & f_3 & \ldots & f_n \end{pmatrix}^* \not\equiv 0.$$

But then also

$$\begin{pmatrix} 1 & D_{1k_1} & D_{2k_2} & \ldots & D_{n-1\,k_{n-1}} \\ f_1 & f_2 & f_3 & \ldots & f_n \end{pmatrix} \not\equiv 0.$$

We have thus the following result[1]

Lemma 1: *Let* $f_1, \ldots, f_n$ *be* n *polynomials in* m *variables that have real coefficients and are independent over the real field. Then there is at least one Wronskian* $\begin{pmatrix} D_1 \ldots D_n \\ f_1 \ldots f_n \end{pmatrix}$ *that does not vanish identically. The same assertion holds if the polynomials have rational coefficients and are independent over the rational field.*

The second assertion of the lemma holds, of course, because, as we saw in §2, the polynomials are also independent over the real field.

---

1. For a detailed study of the generalised Wronskian see, in particular, A. Ostrowski, Math. Z. 4 (1919), 223-230.

6. An identity.

By means of Lemma 1 we shall prove a general identity which is basic for the later induction.

Assume that $m \geq 2$, and denote by

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1} \dots x_m^{i_m} \not\equiv 0$$

a polynomial with integral coefficient. Since

$$A(x_1,\dots,x_m) = \sum_{i_m=0}^{r_m} \left(\sum_{i_1=0}^{r_1} \dots \sum_{i_{m-1}=0}^{r_{m-1}} a_{i_1}\dots{}_{i_{m-1}} x_1^{i_1} \dots x_{m-1}^{i_{m-1}} \right) \cdot x_m^{i_m},$$

it is always possible to write the polynomial in at least one way as a sum

$$A(x_1,\dots,x_m) = \sum_{\nu=1}^{n} P_\nu(x_1,\dots,x_{m-1}) \Sigma_\nu(x_m)$$

where $P_1,\dots,P_n$ are polynomials in $x_1,\dots,x_{m-1}$ and $\Sigma_1,\dots,\Sigma_n$ are polynomials in x_m, all with rational coefficients. From now on choose one such representation for which *the number* n *of terms is a minimum;* then

$$1 \leq n \leq r_m+1.$$

Let us call this the *minimum representation of* A.

In the minimum representation, both the n polynomials

$$P_1(x_1,\dots,x_{m-1}),\dots, P_n(x_1,\dots,x_{m-1})$$

and the n polynomials

$$\Sigma_1(x_m),\dots, \Sigma_n(x_m)$$

are independent over the rational and hence also over the real field (§2). For assume, say, that there are rational numbers $c_1,\dots, c_n$ not all zero such that $c_1 P_1 + \dots + c_n P_n \equiv 0$; let e.g., $c_n \neq 0$. On solving for P_n,

$$P_n \equiv \gamma_1 P_1 + \dots + \gamma_{n-1} P_{n-1}$$

where $\gamma_1,\dots,\gamma_{n-1}$ are rational numbers. Hence we obtain a new representation of A,

$$A(x_1,\dots,x_m) = \sum_{\nu=1}^{n-1} P_{\nu-1}(x_1,\dots,x_{m-1}) \Sigma^*_\nu(x_m) \text{ where } \Sigma^*_\nu(x_m) = \Sigma_\nu(x_m) + \gamma_\nu \Sigma_n(x_m)$$

with at most n-1 terms, contrary to the definition of the minimum representation. The independence of $\Sigma_1,\dots,\Sigma_n$ is proved in the same way.

By Lemma 1 there exist then two Wronskians

$$U^*(x_1,\dots,x_{m-1}) = \begin{pmatrix} D_1^* \dots D_n^* \\ P_1 \dots P_n \end{pmatrix}, \quad V^{**}(x_m) = \begin{pmatrix} D_1^{**} \dots D_n^{**} \\ \Sigma_1 \dots \Sigma_n \end{pmatrix}$$

that do not vanish identically. Here, in the Wronskian U^*,

$$D_\nu^* = \frac{\partial^{j_{\nu 1}+\dots+j_{\nu m-1}}}{\partial x_1^{j_{\nu 1}}\dots\partial x_{m-1}^{j_{\nu m-1}}}$$

with certain non-negative integers $j_{\nu 1},\dots, j_{\nu m-1}$ such that

$$j_{\nu 1}+\dots+j_{\nu m-1} \leq \nu-1 \qquad (\nu = 1,2,\dots, n).$$

On the other hand, in the Wronskian V^{**},

$$D_\nu^{**} = \frac{d^{j_{\nu m}}}{dx_m^{j_{\nu m}}} \quad \text{where} \quad j_{\nu m} = \nu-1 \qquad (\nu = 1,2,\dots, n).$$

Denote by $D_{\mu\nu}$ and $\Delta_{\mu\nu}$ the new operators

$$D_{\mu\nu} = D_\mu^* D_\nu^{**} = \frac{\partial^{j_{\mu 1}+\dots+j_{\mu m-1}+j_{\nu m}}}{\partial x_1^{j_{\mu 1}}\dots\partial x_{m-1}^{j_{\mu m-1}}\partial x_m^{j_{\nu m}}}$$

and

$$\Delta_{\mu\nu} = \frac{1}{j_{\mu 1}!\cdots j_{\mu m-1}!\, j_{\nu m}!} D_{\mu\nu}$$

Further put

$$\overset{*}{W}(x_1,\dots,x_m) = \begin{vmatrix} D_{11}A & D_{12}A & \dots & D_{1n}A \\ D_{21}A & D_{22}A & \dots & D_{2n}A \\ \dots & \dots & \dots & \dots \\ D_{n1}A & D_{n2}A & \dots & D_{nn}A \end{vmatrix},$$

$$W(x_1,\dots,x_m) = \begin{vmatrix} \Delta_{11}A & \Delta_{12}A & \dots & \Delta_{1n}A \\ \Delta_{21}A & \Delta_{22}A & \dots & \Delta_{2n}A \\ \dots & \dots & \dots & \dots \\ \Delta_{n1}A & \Delta_{n2}A & \dots & \Delta_{nn}A \end{vmatrix}.$$

Thus

$$W(x_1,\dots,x_m) = C\, W^*(x_1,\dots,x_m)$$

where $C \neq 0$ is a certain rational number.

On differentiating the minimum representation of A, we obtain the system of identities

$$D_{\lambda\mu} A(x_1,\dots,x_m) = \sum_{\nu=1}^{n} D_\lambda^* P_\nu(x_1,\dots,x_{m-1}).\, D_\mu^{**}\, \Sigma_\nu(x_m) \qquad (\lambda,\mu = 1,2,\dots, n).$$

Therefore, by the multiplication law for determinants,

$$W^*(x_1,\dots,x_m) = U^*(x_1,\dots,x_{m-1})V^{**}(x_m)$$

and hence also

$$W(x_1,\dots,x_m) = CU^*(x_1,\dots, x_{m-1})V^{**}(x_m).$$

It is obvious that all three determinants U^*, V^{**}, and W are polynomials with rational coefficients in some or all of the variables $x_1,\dots, x_m$. Moreover, the stronger result holds that W *has integral coefficients*. For if $j_1,\dots, j_m$ are arbitrary non-negative integers, the partial derivative

$$A_{j_1\dots j_m}(x_1,\dots,x_m) = \frac{\partial^{j_1+\dots+j_m}A(x_1,\dots,x_m)}{j_1!\dots j_m!\ \partial x_1^{j_1}\dots\partial x_m^{j_m}}$$

has the explicit form

$$A_{j_1\dots j_m}(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1}\dots\sum_{i_m=0}^{r_m} a_{i_1\dots i_m}\binom{i_1}{j_1}\dots\binom{i_m}{j_m}x_1^{i_1-j_1}\dots x_m^{i_m-j_m}$$

and hence is a polynomial with integral coefficients. On the other hand, the general element in the determinant W is exactly

$$\Delta_{\mu\nu}A(x_1,\dots,x_m) = A_{j_{\mu1}\dots j_{\mu m-1}j_{\nu m}}(x_1,\dots,x_m),$$

hence is such a polynomial, and so the same is true for W.

Now a well-known theorem due to Gauss states that if f and g are polynomials in any number of variables with rational coefficients such that the product fg has integral coefficients, then there exists a rational number $c \neq 0$ such that both cf and $c^{-1}g$ have integral coefficients. On applying this theorem to the two polynomials CU^* and V^{**}, we find that there are two rational numbers $\mu \neq 0$ and $\nu \neq 0$ such that

$$U(x_1,\dots,x_{m-1}) = uU^*(x_1,\dots,x_{m-1}) \text{ and } V(x_m) = vV^{**}(x_m)$$

have integral coefficients, and that further

$$W(x_1,\dots,x_m) = U(x_1,\dots,x_{m-1})V(x_m).$$

The following result has thus been obtained.

Lemma 2: Let

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1}\dots\sum_{i_m=0}^{r_m} a_{i_1\dots i_m}x_1^{i_1}\dots x_m^{i_m} \not\equiv 0$$

be a polynomial with integral coefficients. There exist a positive integer n *not greater than* r_m+1 *and a system of* n^2 *operators*

$$\Delta_{\mu\nu} = \frac{\partial^{j_{\mu1}+\dots+j_{\mu m-1}+j_{\nu m}}}{j_{\mu1}!\dots j_{\mu m-1}!\,j_{\nu m}!\ \partial x_1^{j_{\mu1}}\dots\partial x_{m-1}^{j_{\mu m-1}}\ \partial x_m^{j_{\nu m}}}$$

where $j_{\mu1},\dots,j_{\mu m-1}, j_{\nu m}$ *are non-negative integers such that*

$$j_{\mu1}+\dots+j_{\mu m-1} \le \mu-1,\ j_{\nu m} = \nu-1 \qquad (\mu,\nu = 1,2,\dots,n),$$

and that the following properties hold. The determinant

$$W(x_1,\dots,x_m) = \begin{vmatrix} \Delta_{11}A & \Delta_{12}A & \dots & \Delta_{1n}A \\ \Delta_{21}A & \Delta_{22}A & \dots & \Delta_{2n}A \\ \vdots & \vdots & & \vdots \\ \Delta_{n1}A & \Delta_{n2}A & \dots & \Delta_{nn}A \end{vmatrix}$$

does not vanish identically, is a polynomial with integral coefficients, and can be written as a product

$$W(x_1,\dots,x_m) = U(x_1,\dots,x_{m-1})V(x_m)$$

where U *and* V *are likewise polynomials with integral coefficients.*

7. Majorants for U, V and W.

If

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1}\dots x_m^{i_m}$$

and

$$B(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} b_{i_1\dots i_m} x_1^{i_1}\dots x_m^{i_m}$$

are two polynomials with real coefficients such that

$$|a_{i_1\dots i_m}| \leq b_{i_1\dots i_m} \quad \text{for all suffixes } i_1,\dots,i_m,$$

then B is said to be a *majorant* or *majoriser* of A, and we write

$$A << B.$$

It is obvious that this relation has the following properties.

If $A << B$ *and* $B << C$, *then* $A << C$.
If $A << B$ *and* $C << D$, *then* $A \mp C << B + D$ *and* $AC << BD$.
If $A << B$, *and* c *is any real number, then* $cA << |c|B$.
The relation $A << B$ *may be differentiated arbitrarily often with respect to any of the variables.*

We also use the notation,

$$\lceil A \rceil = \lceil A(x_1,\dots,x_m) \rceil = \max_{\substack{i_1=0,1,\dots,r_1 \\ \vdots \\ i_m=0,1,\dots,r_m}} |a_{i_1\dots i_m}|,$$

and call $\lceil A \rceil$ the *height* of A. This agrees with the definition of the height of a polynomial in a single variable given in Chapter 3.

We consider now again the polynomial

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1}\dots x_m^{i_m}$$

of Lemma 2 and denote its height by

$$a = \lceil A \rceil.$$

By the binomial theorem,

$$1 + x + \dots + x^r << (1+x)^r.$$

Hence A has the majorant

$$A(x_1,\dots,x_m) << a(1+x_1)^{r_1}\dots(1+x_m)^{r_m}.$$

On differentiating this formula repeatedly, we find that

$$A_{j_1\ldots j_m}(x_1,\ldots,x_m) \ll a\binom{r_1}{j_1}\cdots\binom{r_m}{j_m}(1+x_1)^{r_1-j_1}\cdots(1+x_m)^{r_m-j_m}.$$

Here

$$\binom{r}{j} \leqslant \sum_{j=0}^{r}\binom{r}{j} \leqslant 2^r \quad \text{and} \quad (1+x)^{r-j} \ll (1+x)^r,$$

so that

$$A_{j_1\ldots j_m}(x_1,\ldots,x_m) \ll 2^{r_1+\ldots+r_m}a(1+x_1)^{r_1}\cdots(1+x_m)^{r_m}.$$

In particular, it follows that

$$\Delta_{\mu\nu}A \ll 2^{r_1+\ldots+r_m}a(1+x_1)^{r_1}\cdots(1+x_m)^{r_m} \qquad (\mu,\nu = 1,2,\ldots,n).$$

Now, by its definition as a determinant,

$$W(x_1,\ldots,x_m) = \sum \mp \Delta_{\mu_1 1}A\,\Delta_{\mu_2 2}A\ldots\Delta_{\mu_n n}A$$

where the summation extends over all $n!$ systems of suffixes $\mu_1,\mu_2,\ldots,\mu_n$ that are permutations of $1,2,\ldots,n$. Therefore, on replacing the factors $\Delta_{\mu\nu}A$ by their majorants,

$$W(x_1,\ldots,x_m) \ll n!\,\{2^{r_1+\ldots+r_m}a(1+x_1)^{r_1}\cdots(1+x_m)^{r_m}\}^n.$$

Since

$$(1+x)^{nr} = \sum_{k=0}^{nr}\binom{nr}{k}x^k \ll 2^{nr}(1+x+\ldots+x^{nr}),$$

this formula may be simplified to

$$W(x_1,\ldots,x_m) \ll 2^{2n(r_1+\ldots+r_m)}a^n n!\,(1+x_1+\ldots+x_1^{nr_1})\cdots(1+x_m+\ldots+x_m^{nr_m}).$$

The majorant for W so obtained implies analogous majorants

$$U(x_1,\ldots,x_{m-1}) \ll 2^{2n(r_1+\ldots+r_m)}a^n n!\,(1+x_1+\ldots+x_1^{nr_1})\cdots(1+x_{m-1}+\ldots+x_{m-1}^{nr_{m-1}}),$$

$$V(x_m) \ll 2^{2n(r_1+\ldots+r_m)}a^n n!\,(1+x_m+\ldots+x_m^{nr_m})$$

for U and V. For, by construction, $W=UV$ where U depends only on the variables $x_1,\ldots,x_{m-1}$ and V only on the remaining variable x_m; furthermore, all these polynomials have integral coefficients. Thus the product of any coefficient of U with any coefficient of V is a coefficient of W. Since the non-vanishing coefficients have at least the absolute value 1, it follows that

$$\max(\overline{|U|},\ \overline{|V|}) \leqslant \overline{|W|},$$

whence the asserted majorants for U and V. In this way the following result has been proved.

Lemma 3: *Let* A, U, V, W, *and* n *be as in Lemma* 2, *and let*

$$\alpha = 2^{2n(r_1+\ldots+r_m)}n!\,\overline{|A|}^n;\quad \rho_1 = nr_1,\ldots,\rho_{m-1} = nr_{m-1},\ \rho_m = nr_m.$$

Then

$$\max(\lceil U \rceil, \lceil V \rceil, \lceil W \rceil) \leq \alpha,$$

and the degrees of U in $x_1,\ldots,x_{m-1}$ *do not exceed* $\rho_1,\ldots,\rho_{m-1}$, *the degree of* V *in* x_m *does not exceed* ρ_m, *and the degrees of* W *in* $x_1,\ldots,x_m$ *do not exceed* $\rho_1,\ldots,\rho_m$, *respectively.*

8. The index of a polynomial.

For any $m \geq 1$, let

$$\kappa_1 = \frac{P_1}{Q_1},\ldots,\kappa_m = \frac{P_m}{Q_m}$$

be m rational numbers written in their reduced forms so that

$$(P_1, Q_1) = \ldots = (P_m, Q_m) = 1.$$

The positive integers

$$H_1 = \max(|P_1|, |Q_1|),\ldots, H_m = \max(|P_m|, |Q_m|)$$

are then called the *heights* of $\kappa_1,\ldots,\kappa_m$, respectively. Let further $\rho_1,\ldots,\rho_m$ be m arbitrary positive numbers, and let again

$$A(x_1,\ldots,x_m) = \sum_{i_1=0}^{r_1} \ldots \sum_{i_m=0}^{r_m} a_{i_1\ldots i_m} x_1^{i_1}\ldots x_m^{i_m} \not\equiv 0$$

be a polynomial in $x_1,\ldots,x_m$ with integral coefficients which is not identically zero. Hence the derivatives

$$A_{j_1\ldots j_m}(x_1,\ldots,x_m) = \frac{\partial^{j_1+\ldots+j_m} A(x_1,\ldots,x_m)}{j_1!\ldots j_m!\,\partial x_1^{j_1}\ldots\partial x_m^{j_m}}$$

cannot all vanish at the point $x_1 = \kappa_1,\ldots,x_m = \kappa_m$. Denote by

$$J(A) = J(A; \rho_1,\ldots,\rho_m; \kappa_1,\ldots,\kappa_m)$$

the smallest value of

$$\frac{j_1}{\rho_1} + \ldots + \frac{j_m}{\rho_m}$$

for all systems of suffixes $j_1,\ldots, j_m$ for which

$$A_{j_1\ldots j_m}(\kappa_1,\ldots,\kappa_m) \neq 0,$$

and put $J(A) = \infty$ in the excluded case of the polynomial $A \equiv 0$. The function J(A) of A so defined is called the *index of* A *at the point* $(\kappa_1,\ldots,\kappa_m)$ *relative to* $\rho_1,\ldots,\rho_m$.

This index may also be obtained as follows. By Taylor's formula,

$$A(\kappa_1 + x^{\frac{1}{\rho_1}} x_1,\ldots, \kappa_m + x^{\frac{1}{\rho_m}} x_m) =$$

$$= \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} A_{j_1 \ldots j_m}(\kappa_1, \ldots, \kappa_m) x^{\frac{j_1}{\rho_1} + \ldots + \frac{j_m}{\rho_m}} x_1^{j_1} \ldots x_m^{j_m},$$

where x is a further variable. Hence J(A), for $A \neq 0$, is the exponent of the lowest power of x in this development with a non-zero factor

$$A_{j_1 \ldots j_m}(\kappa_1, \ldots, \kappa_m) x_1^{j_1} \ldots x_m^{j_m}.$$

From this, it follows at once that if $B(x_1, \ldots, x_m)$ is a second polynomial of the same kind, then

(A): $$J(A \mp B) \geq \min\{J(A), J(B)\},$$

(B): $$J(AB) = J(A) + J(B),$$

where the indices are taken at $(\kappa_1, \ldots, \kappa_m)$ relative to $\rho_1, \ldots, \rho_m$. It is further obvious that

$$\text{either } J(A) \geq 0 \text{ or } J(A) = \infty,$$

and that[2]

$$J(A) = 0 \text{ if and only if } A(\kappa_1, \ldots, \kappa_m) \neq 0.$$

We need one further simple property of the index. Let $l_1, \ldots, l_m$ be arbitrary non-negative integers, and let

$$B(x_1, \ldots, x_m) = A_{l_1 \ldots l_m}(x_1, \ldots, x_m).$$

Evidently

$$B_{j_1 \ldots j_m}(\kappa_1, \ldots, \kappa_m) = \binom{j_1^*}{j_1} \cdots \binom{j_m^*}{j_m} A_{j_1^* \ldots j_m^*}(\kappa_1, \ldots, \kappa_m)$$

where

$$j_1^* = j_1 + l_1, \ldots, j_m^* = j_m + l_m \text{ and therefore } \sum_{h=1}^{m} \frac{j_h}{\rho_h} = \sum_{h=1}^{m} \frac{j_h^*}{\rho_h} - \sum_{h=1}^{m} \frac{l_h}{\rho_h}.$$

Since the index cannot be negative, we obtain then the inequality

(C): $$J(A_{l_1 \ldots l_m}) \geq \max\Big(0, J(A) - \sum_{h=1}^{m} \frac{l_h}{\rho_h}\Big).$$

From now on the index J(A) will nearly always be taken relative to $r_1, \ldots, r_m$.

---

2. Put $w(A) = e^{-J(A)}$ if $A \neq 0$, and $W(0) = 0$. The properties of J(A) just stated show that w(A) is a non-archimedean valuation on the ring of polynomials.

9. The upper bound $\Theta_m(a; H_1,\dots, H_m; r_1,\dots, r_m)$

If $H_1,\dots,H_m$ are m positive integers, we denote by $Q(H_1,\dots,H_m)$ the set of all systems of m rational numbers

$$\kappa_1 = \frac{P_1}{Q_1}, \dots, \kappa_m = \frac{P_m}{Q_m}$$

written in their simplified forms,

$$(P_1,Q_1) = (P_2,Q_2) = \dots = (P_m,Q_m) = 1,$$

of heights

$$H_h = \max(|P_h|, |Q_h|) \qquad (h = 1,2,\dots, m).$$

If further $a\geq 1$ is a real number and $r_1,\dots,r_m$ are positive integers, we denote by $R(a; r_1,\dots,r_m)$ the set of all polynomials

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1}\dots x_m^{i_m} \not\equiv 0$$

with integral coefficients which are of height

$$\boxed{A} \leq a.$$

Roth's lemma deals with a number Θ_m defined as follows.

Definition: *Let* $a\geq 1$, *and let* $r_1,\dots,r_m, H_1,\dots,H_m$ *be* 2m *positive integers. The symbol*

$$\Theta_m = \Theta_m(a; r_1,\dots,r_m; H_1,\dots,H_m)$$

denotes the least upper bound of $J(A; r_1,\dots,r_m;\kappa_1,\dots,\kappa_m)$ *extended over all polynomials* $A\in R(A; r_1,\dots,r_m)$ *and over all systems of* m *rational numbers* $(\kappa_1,\dots,\kappa_m)\in Q(H_1,\dots,H_m)$.

Several simple properties of Θ_m follow at once from this definition. First, both sets $R(a; r_1,\dots,r_m)$ and $Q(H_1,\dots,H_m)$ are *finite.* Hence the least upper bound in the definition of Θ_m is *attained,* and *there exist a polynomial* $A\in R$ *and a set of fractions* $(\kappa_1,\dots,\kappa_m)\in Q$ *such that*

$$\Theta_m(a; r_1\dots,r_m; H_1,\dots,H_m) = J(A; r_1,\dots,r_m; \kappa_1,\dots,\kappa_m).$$

Secondly, the set $R(a; r_1,\dots,r_m)$ does not lose elements when a increases; hence Θ_m *is a non-decreasing function of* a. Third, as we are considering now indices relative to $r_1,\dots,r_m$, $J(A)$ is equal to a sum

$$\sum_{h=1}^{m} \frac{j_h}{r_h} \quad \text{where} \quad 0 \leq j_1 \leq r_1,\dots,0 \leq j_m \leq r_m,$$

and it follows at once that $0 \leq J(A) \leq m$ and hence that

$$0 \leq \Theta_m \leq m.$$

10. An upper bound for $\Theta_1(a; r; H)$.

We begin with the study of Θ_1. Let

$$A(x) = \sum_{i=0}^{r} a_i x^i \neq 0$$

be a polynomial in one variable, with integral coefficients such that

$$\boxed{A} = \max(|a_0|, |a_1|, \ldots, |a_r|) \leq a.$$

Let the rational number $\kappa = \frac{P}{Q}$ be written in its simplest form, thus $(P,Q)=1$, and assume that

$$H = \max(|P|, |Q|) \geq 2.$$

Hence $\kappa \neq 0$ and $P \neq 0$, $Q \neq 0$.

Suppose $J(A; r; \kappa) > 0$. Then κ is a zero of $A(x)$, say of the exact order $j > 0$, and $A(x)$ is divisible by $(x-\kappa)^j$. By Gauss's lemma from §6, $A(x)$ can then be written as

$$A(x) = (Qx-P)^j B(x)$$

where $B(x)$ is a certain polynomial with integral coefficients. Hence the lowest and the highest non-zero coefficients of $A(x)$ are divisible by P^j and Q^j, respectively. It follows that

$$H^j \leq \boxed{A} \leq a,$$

hence that

$$J(A; r; \kappa) = \frac{j}{r} \leq \frac{\log a}{r \log H},$$

a result valid also when $J(A; r; \kappa) = 0$. Therefore always

$$(\Theta_1): \qquad \Theta_1(a; r; H) \leq \frac{\log a}{r \log H} \quad \text{if} \quad H \geq 2.$$

11. The property Γ_M.

The induction proof of Roth's lemma in the next sections becomes simpler if the following notation is used.

Denote by t a constant such that

$$0 < t \leq 1.$$

If $b \geq 1$ is a real number, and $s_1, \ldots, s_M$, $K_1, \ldots, K_M$ are positive integers, *the ordered system of numbers*

$$b,\ s_1, \ldots, s_M,\ K_1, \ldots, K_M$$

is said to have the property Γ_M *if either,* (i)

$$M = 1,\ K_1 \geq 2,\ b \leq K_1^{s_1 t};$$

or, (ii), *simultaneously* $M \geq 2$ *and*

$$\max\left(\frac{s_2}{s_1}, \frac{s_3}{s_2}, \ldots, \frac{s_M}{s_{M-1}}\right) \leq t;$$

$$s_h \log K_h \geq s_1 \log K_1 \qquad (h = 1,2,\ldots, M);$$

$$K_1 \geq 2^{\frac{1}{t}(M-1)M(2M+1)};$$

$$b \leq K_1^{\frac{1}{M} s_1 t}.$$

Therefore the following inequalities also hold,

$$s_1 \geq s_2 \geq \ldots \geq s_M; \ \sum_{h=1}^{M} s_h \leq M s_1; \ K_1 \geq 2,\ K_2 \geq 2, \ldots, K_M \geq 2.$$

By way of application, let $m \geq 2$; let the ordered system of numbers

$$a, r_1, \ldots, r_m, H_1, \ldots, H_m$$

have the property Γ_m; let n be an integer such that

$$1 \leq n \leq r_m + 1;$$

and let

$$b = 2^{2n(r_1+\ldots+r_m)} n!\, a^n; \quad \rho_1 = nr_1,\ \rho_2 = nr_2, \ldots, \rho_{m-1} = nr_{m-1}.$$

Then the new ordered system of numbers

$$b, \rho_1, \ldots, \rho_{m-1}, H_1, \ldots, H_{m-1}$$

has the property Γ_{m-1}.

Proof: The first inequalities

$$\max\left(\frac{\rho_2}{\rho_1}, \frac{\rho_3}{\rho_2}, \ldots, \frac{\rho_{m-1}}{\rho_{m-2}}\right) \leq t,$$

$$\rho_h \log H_h \geq \rho_1 \log H_1 \qquad (h = 1,2,\ldots, m-1)$$

are for $m \geq 3$ immediate consequences of the assumption that

$$\max\left(\frac{r_2}{r_1}, \frac{r_3}{r_2}, \ldots, \frac{r_m}{r_{m-1}}\right) \leq t;$$

$$r_h \log H_h \geq r_1 \log H_1 \qquad (h = 1,2,\ldots, m).$$

Next, by hypothesis,

$$H_1 \geq 2^{\frac{1}{t}(m-1)m(2m+1)},$$

whence, trivially, also

$$H_1 \geq 2^{\frac{1}{t}(m-2)(m-1)(2m-1)}.$$

Finally,

$$n \leqslant r_m + 1 \leqslant r_1 + 1 \leqslant 2^{r_1}, \quad \text{hence} \quad n! \leqslant n^n \leqslant 2^{nr_1} = 2^{\rho_1},$$

and, by assumption,

$$a \leqslant H_1^{\frac{1}{m}r_1 t}, \quad r_1 + r_2 + \cdots + r_m \leqslant mr_1.$$

Therefore,

$$b \leqslant 2^{2n \cdot mr_1} \cdot 2^{\rho_1} \cdot a^n \leqslant 2^{(2m+1)\rho_1} H_1^{\frac{1}{m}\rho_1 t} \leqslant H_1^{\frac{1}{m-1}\rho_1 t},$$

because

$$H_1^{\frac{1}{m-1}\rho_1 t} \Big/ H_1^{\frac{1}{m}\rho_1 t} = H_1^{\frac{1}{m(m-1)}\rho_1 t} \geqslant 2^{(2m+1)\rho_1}.$$

12. A recursive inequality for Θ_m. I.

Let again t be a constant such that

$$0 < t \leqslant 1.$$

We assume that $m \geqslant 2$ and that the ordered system of numbers

$$a, r_1, \ldots, r_m, H_1, \ldots, H_m$$

has the property Γ_m.

It was shown in §9 that there exist a polynomial

$$A(x_1, \ldots, x_m) \in R(a; r_1, \ldots, r_m)$$

and a set of fractions

$$(\kappa_1, \ldots, \kappa_m) \in Q(H_1, \ldots, H_m)$$

such that

$$\Theta_m(a; r_1, \ldots, r_m; H_1, \ldots, H_m) = J(A; r_1, \ldots, r_m; \kappa_1, \ldots, \kappa_m).$$

Denote by n the integer with $1 \leqslant n \leqslant r_m + 1$, and by $U(x_1, \ldots, x_{m-1})$, $V(x_m)$, and $W(x_1, \ldots, x_m)$ the three polynomials that correspond to A by Lemmas 2 and 3, and put again

$$b = 2^{2n(r_1 + \cdots + r_m)} n!\, a^n; \quad \rho_1 = nr_1, \ldots, \rho_{m-1} = nr_{m-1}, \rho_m = nr_m.$$

As has just been proved, the ordered system of numbers

$$b, \rho_1, \ldots, \rho_{m-1}, H_1, \ldots, H_{m-1}$$

has then the property Γ_{m-1}.

From the construction and from Lemma 3,

$$\overline{A} \leqslant a; \quad \overline{U} \leqslant \alpha, \quad \overline{V} \leqslant \alpha, \quad \overline{W} \leqslant \alpha,$$

where

$$\alpha = 2^{2n(r_1 + \cdots + r_m)} n!\, \overline{A}^n \leqslant b.$$

Therefore the upper bounds for the degrees of U, V, and W imply that

$$U(x_1,\dots,x_{m-1})\in R(b;\rho_1,\dots,\rho_{m-1}),\ V(x_m)\in R(b;\rho_m),$$

$$W(x_1,\dots,x_m)\in R(b;\rho_1,\dots,\rho_m).$$

Hence, in particular, with the same fractions $\kappa_1,\dots,\kappa_{m-1},\kappa_m$ as above,

$$J(U;\rho_1,\dots,\rho_{m-1};\kappa_1,\dots,\kappa_{m-1}) \leqslant \Theta_{m-1}(b;\rho_1,\dots,\rho_{m-1};H_1,\dots,H_{m-1}),$$

$$J(V;\rho_m;\kappa_m) \leqslant \Theta_1(b;\rho_m;H_m).$$

From the identity

$$W(x_1,\dots,x_m) = U(x_1,\dots,x_{m-1})V(x_m)$$

and from the multiplicative property (B) of the index, it follows that

$$J(W;\rho_1,\dots,\rho_m;\ _1,\dots,\kappa_m) = J(U;\rho_1,\dots,\rho_{m-1};\kappa_1,\dots,\kappa_{m-1}) + J(V;\rho_m;\kappa_m),$$

or

$$J(W;\rho_1,\dots,\rho_m;\kappa_1,\dots,\kappa_m) \leqslant \Phi_m,$$

where, for shortness,

$$\Phi_m = \Theta_{m-1}(b;\rho_1,\dots,\rho_{m-1};H_1,\dots,H_{m-1}) + \Theta_1(b;\rho_m;H_m).$$

Instead, we may also write

$$J(W;r_1,\dots,r_m;\kappa_1,\dots,\kappa_m) \leqslant n\,\Phi_m,$$

because $\rho_h = nr_h$ for all h, and so, by the definition of the index,

$$J(W;r_1,\dots,r_m;\kappa_1,\dots,\kappa_m) = n\,J(W;\rho_1,\dots,\rho_m;\kappa_1,\dots,\kappa_m).$$

Since from now on only indices of polynomials at the fixed point $(\kappa_1,\dots,\kappa_m)$ relative to the fixed integers $r_1,\dots,r_m$ will occur, we shall write for these indices simply J(W), J(A), etc.

13. A recursive inequality for Θ_m. II.

In the inequality

$$J(W) \leqslant n\Phi_m$$

just proved, we can give a lower bound for J(W) in terms of J(A).

For, as in §7,

$$W(x_1,\dots,x_m) = \sum \mp \Delta_{\mu_1 1}A\,\Delta_{\mu_2 2}A\dots\Delta_{\mu_n n}A,$$

where the systems of suffixes $\mu_1,\dots,\mu_n$ run over all n! permutations of 1,2,...,n, while the operators $\Delta_{\mu\nu}$ are of the form

$$\Delta_{\mu\nu} = \frac{\partial^{j_{\mu 1}+\dots+j_{\mu m-1}+j_{\nu m}}}{j_{\mu 1}!\dots j_{\mu m-1}!\,j_{\nu m}!\,\partial x_1^{j_{\mu 1}}\dots\partial x_{m-1}^{j_{\mu m-1}}\,\partial x_m^{j_{\nu m}}},$$

and the j's are non-negative integers such that

$$j_{\mu 1}+\dots+j_{\mu m-1} \leqslant \mu-1, \qquad j_{\nu m} = \nu-1 \qquad (\mu,\nu = 1,2,\dots,n).$$

Therefore

$$\Delta_{\mu\nu}A = A_{j_{\mu 1}\dots j_{\mu m-1}j_{\nu m}},$$

so that property (C) of the index implies the inequality

$$J(\Delta_{\mu\nu}A) \geqslant \max\left(0,\, J(A) - \sum_{h=1}^{m-1} \frac{j_{\mu h}}{r_h} - \frac{j_{\nu m}}{r_m}\right).$$

But

$$r_1 \geqslant r_2 \geqslant \dots \geqslant r_{m-1} \geqslant r_m;\ \mu-1 \leqslant n-1 \leqslant r_m \leqslant tr_{m-1},$$

and so

$$J(\Delta_{\mu\nu}A) \geqslant \max\left(0,\, J(A) - \frac{1}{r_{m-1}} \sum_{h=1}^{m-1} j_{\mu h} - \frac{j_{\nu m}}{r_m}\right) \geqslant$$

$$\geqslant \max\left(0,\, J(A) - \frac{\mu-1}{r_{m-1}} - \frac{\nu-1}{r_m}\right) \geqslant \max\left(0,\, J(A)-t-\frac{\nu-1}{r_m}\right).$$

Therefore, finally, by the properties (A) and (B) of the index,

$$J(W) \geqslant \sum_{\nu=1}^{n} \max\left(0,\, J(A)-t-\frac{\nu-1}{r_m}\right).$$

This inequality can be simplified. For shortness put

$$N = [\{J(A)-t\}r_m] + 1,$$

where $[x]$ denotes as usual the integral part of x. Hence

$$N-1 \leqslant \{J(A)-t\}r_m < N.$$

We shall now distinguish the two cases $n \leqslant N$ and $n > N$.

The case $n \leqslant N$. Evidently

$$n-1 \leqslant N-1 \leqslant \{J(A)-t\}r_m$$

and hence

$$\frac{\nu-1}{r_m} \leqslant \frac{n-1}{r_m} \leqslant J(A)-t \qquad (\nu = 1,2,\dots,n).$$

Therefore

$$J(W) \geqslant \sum_{\nu=1}^{n} \{J(A)-t-\frac{\nu-1}{r_m}\} = n\{J(A)-t\} - \frac{n(n-1)}{2r_m} \geqslant$$

$$\geqslant n\{J(A)-t\} - \frac{n}{2}\{J(A)-t\} = \frac{n}{2}\{J(A)-t\},$$

whence

(1): $$J(A) \leqslant t + \frac{2}{n} J(W) \qquad \text{if } n \leqslant N.$$

The case $n>N$. Assume, for the moment, that

$$J(A) \geqslant t \quad \text{and thus} \quad N \geqslant 1.$$

Then

$$\max\left(0, J(A)-t-\frac{\nu-1}{r_m}\right) = \begin{cases} J(A)-t-\frac{\nu-1}{r_m} & \text{if } 1 \leqslant \nu \leqslant N, \\ 0 & \text{if } N+1 \leqslant \nu \leqslant n, \end{cases}$$

and it follows that

$$J(W) \geqslant \sum_{\nu=1}^{N} \left\{J(A)-t-\frac{\nu-1}{r_m}\right\} = N\{J(A)-t\} - \frac{N(N-1)}{2r_m} \geqslant$$

$$\geqslant N\{J(A)-t\} - \frac{N}{2}\{J(A)-t\} = \frac{N}{2}\{J(A)-t\} > \frac{r_m}{2}\{J(A)-t\}^2.$$

On solving this inequality for $J(A)$, we find that

(2): $$J(A) \leqslant t + \sqrt{\frac{2}{r_m} J(W)} \leqslant t + 2\sqrt{\frac{1}{n} J(W)} \quad \text{if} \quad n > N,$$

because

$$n \leqslant r_m + 1, \; \frac{r_m}{2} \geqslant \frac{r_m}{r_m+1} \cdot \frac{n}{2} \geqslant \frac{1}{2} \cdot \frac{n}{2} = \frac{n}{4}.$$

In the case $J(A)<t$ so far excluded this estimate is still valid.

The two inequalities (1) and (2) show that in both cases $n \leqslant N$ and $n>N$ the same estimate

$$J(A) \leqslant t + 2 \max\left(\frac{1}{n} J(W), \sqrt{\frac{1}{n} J(W)}\right)$$

holds. Since $J(W) \leqslant n\Phi_m$, it follows that always

$$J(A) \leqslant t + 2 \max(\Phi_m, \sqrt{\Phi_m}).$$

In this inequality, A was chosen such that

$$J(A) = J(A; r_1, \ldots, r_m; \kappa_1, \ldots, \kappa_m) = \Theta_m(a; r_1, \ldots, r_m; H_1, \ldots, H_m),$$

and Φ_m had the value

$$\Phi_m = \Theta_{m-1}(b; \rho_1, \ldots, \rho_{m-1}; H_1, \ldots, H_{m-1}) + \Theta_1(b; \rho_m; H_m).$$

We apply now the inequality (Θ_1) of §10 which shows that

$$\Theta_1(b; \rho_m; H_m) \leqslant \frac{\log b}{\rho_m \log H_m}.$$

By definition, $\rho_1 = nr_1$ and $\rho_m = nr_m$, and also

$$r_m \log H_m \geqslant r_1 \log H_1, \quad \text{hence} \quad \rho_m \log H_m \geqslant \rho_1 \log H_1.$$

It was further shown in §11 that

$$b \leqslant H_1^{\frac{1}{m-1}\rho_1 t}.$$

Therefore

$$\Theta_1(b;\rho_m;H_m) \leqslant \frac{\log b}{\rho_1 \log H_1} \leqslant \frac{t}{m-1} \leqslant t$$

because $m \geqslant 2$.

The long proof of the last sections has thus lead to the following recursive inequality.

Lemma 4: *Let* $m \geqslant 2$, *and let*

$$a; r_1, \ldots, r_m; H_1, \ldots, H_m$$

be any ordered system of numbers with the property Γ_m. *Then there exists an ordered system of numbers*

$$b; \rho_1, \ldots, \rho_{m-1}; H_1, \ldots, H_{m-1}$$

with the property Γ_{m-1} *such that*

$$\Theta_m(a; r_1, \ldots, r_m; H_1, \ldots, H_m) \leqslant t + 2\max(\Psi_m, \sqrt{\Psi_m})$$

where

$$\Psi_m = \Theta_{m-1}(b; \rho_1, \ldots, \rho_{m-1}; H_1, \ldots, H_{m-1}) + t.$$

14. Proof of Roth's Lemma.

It is now easy to prove

Roth's Lemma: *Put* $c_m = 2^{m+1} - 3$. *If the ordered system of numbers*

$$a, r_1, \ldots, r_m, H_1, \ldots, H_m$$

has the property Γ_m, *then*

$$\Theta_m(a; r_1, \ldots, r_m; H_1, \ldots, H_m) \leqslant c_m t^{2^{-(m-1)}}.$$

Proof: We procede by induction for m. First let $m=1$, hence $H_1 \geqslant 2$ and $a \leqslant H_1^{r_1 t}$. The estimate (Θ_1) of §10 implies then that

$$\Theta_1(a; r_1; H_1) \leqslant \frac{\log a}{r_1 \log H_1} \leqslant t = c_1 t,$$

as asserted. Secondly assume that $m \geqslant 2$, and that the assertion has already been proved for all ordered systems of numbers

$$b, \rho_1, \ldots, \rho_{m-1}, H_1, \ldots, H_{m-1}$$

with the property Γ_{m-1}; it suffices to prove that it then is true also for all ordered systems of numbers

$$a, r_1, \ldots, r_m, H_1, \ldots, H_m$$

with the property Γ_m. By this induction hypothesis, the expression Ψ_m in Lemma 4 satisfies the inequality

$$\Psi_m = \Theta_{m-1}(b; \rho_1, \ldots, \rho_{m-1}; H_1, \ldots, H_{m-1}) + t \leqslant c_{m-1} t^{2^{-(m-2)}} + t,$$

and therefore Lemma 4 implies that

$$\Theta_m(a; r_1,\dots,r_m; H_1,\dots,H_m) \leq t + 2\max\left(c_{m-1}t^{2^{-(m-2)}} + t, \sqrt{c_{m-1}t^{2^{-(m-2)}} + t}\right).$$

Now $0 < t \leq 1$ and $c_{m-1} \geq 1$. The expression

$$t + 2\max\left(c_{m-1}t^{2^{-(m-2)}} + t, \sqrt{c_{m-1}t^{2^{-(m-2)}} + t}\right)$$

of this inequality is therefore certainly not larger than

$$t^{2^{-(m-1)}} + 2\max\left(c_{m-1}t^{2^{-(m-1)}} + t^{2^{-(m-1)}}, \sqrt{(c_{m-1}^2 + 2c_{m-1})t^{2^{-(m-2)}} + t^{2^{-(m-2)}}}\right) =$$

$$= (2c_{m-1} + 3)t^{2^{-(m-1)}} = c_m t^{2^{-(m-1)}},$$

whence the assertion.

We conclude this chapter by stating Roth's Lemma in a slightly weaker, but more convenient explicit form, as follows.

Theorem 1: *Let* $0<t\leq 1$. *Let* $a\geq 1$ *be a real number, and let* $r_1,\dots,r_m$, $H_1,\dots,H_m$, *where* $m\geq 2$, *be positive integers such that*

$$r_{h+1} \leq r_h t \qquad (h = 1,2,\dots,m-1),$$

$$r_h \log H_h \geq r_1 \log H_1 \qquad (h = 2,3,\dots,m),$$

$$H_1 \geq 2^{\frac{1}{t}(m-1)m(2m+1)},$$

$$a \leq H_1^{\frac{1}{m} r_1 t}.$$

Let

$$\kappa_1 = \frac{P_1}{Q_1}, \dots, \kappa_m = \frac{P_m}{Q_m}$$

be rational numbers such that

$$(P_h, Q_h) = 1, \ \max(|P_h|, |Q_h|) = H_h \qquad (h = 1,2,\dots,m).$$

Finally let $A(x_1,\dots,x_m)$ *be a polynomial of the form*

$$A(x_1,\dots,x_m) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} a_{i_1\dots i_m} x_1^{i_1} \dots x_m^{i_m}$$

which is not identically zero and has integral coefficients such that

$$|a_{i_1\dots i_m}| \leq a \text{ for all } i_1,\dots, i_m.$$

Then there exist non-negative integers $j_1,\dots,j_m$ *satisfying*

$$\sum_{h=1}^{m} \frac{j_h}{r_h} \leq 2^{m+1} t^{2^{-(m-1)}}$$

such that

$$A_{j_1\dots j_m}(\kappa_1,\dots,\kappa_m) \neq 0.$$