MATHEMATIKA

A JOURNAL OF PURE AND APPLIED MATHEMATICS

VOL. 2. PART 1. JUNE, 1955. No. 3.

RATIONAL APPROXIMATIONS TO ALGEBRAIC NUMBERS

K. F. ROTH

1. It was remarked by Liouville in 1844 that there is an obvious limit to the accuracy with which algebraic numbers can be approximated by rational numbers; if α is an algebraic number of degree n (at least 2) then†

$$\left|\alpha - \frac{h}{q}\right| > \frac{A}{q^n}$$

for all rational numbers h/q, where A is a positive number depending only on α.

More precise and very much more profound results were proved by Thue in 1908, by Siegel in 1921, and by Dyson in 1947. Suppose the inequality

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^\kappa} \tag{1}$$

is satisfied by infinitely many rational numbers h/q. Then Thue proved that $\kappa \leqslant \frac{1}{2}n+1$, Siegel proved that

$$\kappa \leqslant s + \frac{n}{s+1} \quad \text{for } s = 1, 2, \ldots, n-1,$$

and Dyson‡ proved that $\kappa \leqslant \sqrt{(2n)}$.

It was conjectured by Siegel that in reality $\kappa \leqslant 2$, and it is the purpose of this paper to prove that conjecture. Our result is accordingly as follows.

† The result is an immediate deduction from the definition of an algebraic number; see, for example, Davenport, *The Higher Arithmetic* (London 1952), 165–167.

‡ *Acta Mathematica*, 79 (1947), 225–240. The algebraic part of Dyson's work was simplified by Mahler, *Proc. K. Akad. Wet. Amsterdam*, 52 (1949), 1175–1184. Another proof of Dyson's result was given by Schneider in *Archiv der Math.*, 1 (1948–9), 288–295. Dyson's result (with a generalization) was apparently obtained independently by Gelfond; see his *Transcendental and algebraic numbers* (Moscow 1952, in Russian), Chapter 1.

Theorem. *Let α be any algebraic number, not rational. If* (1) *has an infinity of solutions in integers h and q* ($q > 0$) *then* $\kappa \leqslant 2$.

The inequality $\kappa \leqslant 2$ is, of course, the best possible, since every irrational number, whether algebraic or not, has infinitely many rational approximations satisfying (1) with $\kappa = 2$.

The above theorem, like its predecessors, has applications to other arithmetical questions, and in particular to the theory of Diophantine equations†. Suppose $f(x, y)$ is a homogeneous irreducible polynomial of degree n with integral coefficients. It follows easily from the theorem that if the inequality

$$|f(x, y)| < (|x|+|y|)^{n-\kappa}$$

has an infinity of solutions in integers x, y then $\kappa \leqslant 2$. Thus if $g(x, y)$ is any polynomial, not necessarily homogeneous, every term in which has total degree at most $n-3$, then the equation

$$f(x, y) = g(x, y)$$

can have only a finite number of integral solutions.

Various generalizations and analogues of the Thue-Siegel-Dyson theorem are known, and it seems probable that the method of the present paper will lead to improvements in many such results. Certainly this is the case for all the results in Siegel's basic memoir‡, one of the most important of which concerns approximation to an algebraic number by algebraic numbers of given degree. Improving this by the method of the present paper, I have obtained the following generalization of the theorem stated above. *Let α be any algebraic number, not rational. If the inequality*

$$|\alpha-\beta| < \big(H(\beta)\big)^{-\kappa}$$

is satisfied by infinitely many algebraic numbers β of degree g, then $\kappa \leqslant 2g$. Here $H(\beta)$ denotes the maximum absolute value of the rational integral coefficients in the primitive irreducible equation satisfied by β.

As regards the substance of the present paper, it will be appreciated that many of the ideas and methods used are not new. The novel part of the proof is that culminating in Lemma 7, and even here we make much use of ideas that have occurred before in the literature of the subject.

I am greatly indebted to Prof. Davenport for his constant encouragement while I was working on the problem, and for rewriting my original manuscript in a form suitable for publication. In particular, my original

† See Skolem, *Diophantische Gleichungen* (Ergebnisse der Math. V4, Berlin, 1938), Chapter 6, §2.

‡ *Math. Zeitschrift*, 9 (1921), 173–213.

manuscript referred to a lemma of Schneider†, and this Prof. Davenport replaced by the much simpler Lemma 8. This is one of many simplifications and improvements he has introduced.

2. If $\phi_0(x)$, $\phi_1(x)$, ..., $\phi_{l-1}(x)$ are l polynomials in a single variable, the determinant

$$\det\left(\frac{1}{\mu!}\frac{d^\mu}{dx^\mu}\phi_\nu(x)\right) \quad (\mu,\ \nu = 0,\ 1,\ \ldots,\ l-1)$$

is called their Wronskian. If $\phi_0(x)$, ..., $\phi_{l-1}(x)$ have rational coefficients, it is well known that their Wronskian vanishes identically if and only if they are linearly dependent, that is, satisfy identically a linear relation

$$c_0\phi_0(x)+\ldots+c_{l-1}\phi_{l-1}(x) = 0$$

with rational constant coefficients c_0, ..., c_{l-1}.

We define generalized Wronskians‡ for polynomials in p variables as follows. We consider differential operators of the form

$$\Delta = \frac{1}{i_1!\ldots i_p!}\left(\frac{\partial}{\partial x_1}\right)^{i_1}\ldots\left(\frac{\partial}{\partial x_p}\right)^{i_p}, \tag{2}$$

and we call $i_1+\ldots+i_p$ the *order* of the operator Δ. If

$$\phi_0(x_1,\ \ldots,\ x_p),\ \ldots,\ \phi_{l-1}(x_1,\ \ldots,\ x_p)$$

are l polynomials in p variables, and

$$\Delta_0,\ \Delta_1,\ \ldots,\ \Delta_{l-1}$$

are any differential operators of the form (2) whose orders are at most 0, 1, ..., $l-1$ respectively, we call the determinant

$$G(x_1,\ \ldots,\ x_p) = \det\left(\Delta_\mu\phi_\nu(x_1,\ \ldots,\ x_p)\right) \quad (\mu,\ \nu = 0,\ \ldots,\ l-1)$$

a generalized Wronskian of ϕ_0, ..., ϕ_{l-1}. If $p>1$ and $l>1$ there is more than one such generalized Wronskian. It is plain that if ϕ_0, ..., ϕ_{l-1} are linearly dependent then all their generalized Wronskians vanish identically. We proceed to prove the converse§.

† *J. für die reine und angew. Math.*, 175 (1936), 182–192, Lemma 1, formula (7). This paper contains a proof that $\kappa \leqslant 2$ provided that the solutions of (1) satisfy a certain very restrictive condition.

‡ Since writing this paper I find that generalized Wronskians were used by Siegel [*Math. Annalen*, 84 (1921), 80–99] in a similar connection. See also Kellogg, *Comptes rendus des séances de la Soc. Math. de France*, 41 (1912), 19–21, where the main result (Lemma 1 below) is stated without proof.

§ It should perhaps be remarked (though it is immaterial to our argument) that the generalized Wronskians and their derivatives may satisfy identities, by virtue of which the vanishing of some of the generalized Wronskians implies the vanishing of the others, Lemma 1. *If $\phi_0(x_1, \ldots, x_p), \ldots, \phi_{l-1}(x_1, \ldots, x_p)$ are l linearly independent polynomials in p variables, with rational coefficients, then at least one of their generalized Wronskians does not vanish identically.*

Proof. Let k be an integer which is greater than the degrees of all the polynomials $\phi_0, \ldots, \phi_{l-1}$ in each of the separate variables $x_1, \ldots, x_p$. Consider the l polynomials

$$\phi_\nu(t, t^k, t^{k^2}, \ldots, t^{k^{p-1}}) \quad (\nu = 0, \ldots, l-1) \tag{3}$$

in the single variable t. These polynomials are linearly independent. For let

$$\phi_\nu(x_1, \ldots, x_p) = \sum_{s_1=0}^{k-1} \ldots \sum_{s_p=0}^{k-1} b^{(\nu)}(s_1, \ldots, s_p)\, x_1^{s_1} \ldots x_p^{s_p};$$

if the polynomials (3) were linearly dependent there would be an identity in t of the form

$$\sum_{\nu=0}^{l-1} c_\nu \sum_{s_1=0}^{k-1} \ldots \sum_{s_p=0}^{k-1} b^{(\nu)}(s_1, \ldots, s_p)\, t^{s_1+ks_2+\ldots+k^{p-1}s_p} = 0.$$

Since the representation of an integer in the form

$$s_1+ks_2+\ldots+k^{p-1}s_p \quad (0 \leqslant s_1 \leqslant k-1, \ldots, 0 \leqslant s_p \leqslant k-1)$$

is unique, this identity would imply the corresponding identity

$$\sum_{\nu=0}^{l-1} c_\nu \phi_\nu(x_1, \ldots, x_p) = 0.$$

It follows that the Wronskian of the l polynomials (3), namely

$$W(t) = \det\left(\frac{1}{\mu!}\left(\frac{d}{dt}\right)^{\mu} \phi_\nu(t, t^k, \ldots, t^{k^{p-1}})\right) \quad (\mu, \nu = 0, \ldots, l-1), \tag{4}$$

does not vanish identically. Now

$$\frac{d}{dt} = \frac{\partial}{\partial x_1} + kt^{k-1}\frac{\partial}{\partial x_2} + \ldots + k^{p-1}t^{k^{p-1}-1}\frac{\partial}{\partial x_p},$$

where the operators on the right are applied to a polynomial in $x_1, \ldots, x_p$ and these variables are subsequently replaced by $t, \ldots, t^{k^{p-1}}$. By induction on μ, we see that the operator $(d/dt)^\mu$ is expressible as a linear combination of differential operators on $x_1, \ldots, x_p$ of the form (2), of orders not exceeding μ:

$$\left(\frac{d}{dt}\right)^{\mu} = f_1(t)\,\Delta^{(1)} + \ldots + f_r(t)\,\Delta^{(r)},$$

where r depends only on μ and p, and $\Delta^{(1)}, \ldots, \Delta^{(r)}$ are operators of orders not exceeding μ, and $f_1(t), \ldots, f_r(t)$ are polynomials with rational coefficients. Substituting in (4) and expressing the determinant as a sum of other determinants, we obtain an expression for W of the form

$$W(t) = g_1(t)\, G^{(1)}(t, \ldots, t^{k^{p-1}}) + \ldots + g_s(t)\, G^{(s)}(t, \ldots, t^{k^{p-1}}),$$

where $G^{(1)}, \ldots, G^{(s)}$ are certain generalized Wronskians of $\phi_0, \ldots, \phi_{l-1}$ and $g_1(t), \ldots, g_s(t)$ are polynomials in t.

Since $W(t)$ does not vanish identically, there is some i for which $G^{(i)}(t, t^k, \ldots, t^{k^{p-1}})$ does not vanish identically, and *a fortiori* $G^{(i)}(x_1, \ldots, x_p)$ does not vanish identically.

3. LEMMA 2. *Let $R(x_1, \ldots, x_p)$ be a polynomial in $p \geqslant 2$ variables, with integral coefficients, which is not identically zero. Let R be of degree at most r_j in x_j for $j = 1, \ldots, p$. Then there exists an integer l satisfying*

$$1 \leqslant l \leqslant r_p + 1, \tag{5}$$

and there exist differential operators $\Delta_0, \ldots, \Delta_{l-1}$ on the variables $x_1, \ldots, x_{p-1}$, of orders at most $0, \ldots, l-1$ respectively, such that if

$$F(x_1, \ldots, x_p) = \det\left(\Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial x_p}\right)^\nu R\right) \quad (\mu, \nu = 0, \ldots, l-1) \tag{6}$$

then

(i) *F has integral coefficients and is not identically zero;*

(ii) *we have*

$$F(x_1, \ldots, x_p) = U(x_1, \ldots, x_{p-1})\, V(x_p), \tag{7}$$

where U and V have integral coefficients, and U is of degree at most lr_j in x_j for $j = 1, \ldots, p-1$ and V is of degree at most lr_p in x_p.

Proof. We consider all representations of R in the form

$$R(x_1, \ldots, x_p) = \phi_0(x_p)\, \psi_0(x_1, \ldots, x_{p-1}) + \ldots + \phi_{l-1}(x_p)\, \psi_{l-1}(x_1, \ldots, x_{p-1}),$$

where the ϕ_ν and ψ_ν are polynomials with *rational* coefficients, subject to the condition that the ϕ_ν are of degree at most r_p and the ψ_ν of degree at most r_j in x_j for $j = 1, \ldots, p-1$. Such a representation is possible, *e.g.* with $l-1 = r_p$ and $\phi_\nu(x_p) = x_p{}^\nu$. From all such representations we select one for which l is least. Then

$$\phi_0(x_p), \ldots, \phi_{l-1}(x_p)$$

are linearly independent. For if not, say

$$\phi_{l-1} = d_0 \phi_0 + \cdots + d_{l-2} \phi_{l-2}$$

with rational coefficients $d_0, \ldots, d_{l-2}$, we should have

$$R = \phi_0(\psi_0 + d_0 \psi_{l-1}) + \cdots + \phi_{l-2}(\psi_{l-2} + d_{l-2} \psi_{l-1}),$$

contrary to the definition of l. Similarly

$$\psi_0(x_1, \ldots, x_{p-1}), \ldots, \psi_{l-1}(x_1, \ldots, x_{p-1})$$

are linearly independent. Also $1 \leqslant l \leqslant r_p+1$.

Let $W(x_p)$ denote the Wronskian of $\phi_0(x_p), \ldots, \phi_{l-1}(x_p)$, so that W is a polynomial with rational coefficients, not identically zero. Let $G(x_1, \ldots, x_{p-1})$ denote some generalized Wronskian of

$$\psi_0(x_1, \ldots, x_{p-1}), \ldots, \psi_{l-1}(x_1, \ldots, x_{p-1})$$

which is not identically zero, the existence of such a generalized Wronskian being assured by Lemma 1. Then

$$W(x_p) = \det\left(\frac{1}{\mu\,!}\left(\frac{d}{dx_p}\right)^{\mu} \phi_\nu(x_p)\right) \quad (\mu, \nu = 0, \ldots, l-1)$$

and

$$G(x_1, \ldots, x_{p-1}) = \det\left(\Delta_\mu \psi_\nu(x_1, \ldots, x_{p-1})\right) \quad (\mu, \nu = 0, \ldots, l-1),$$

where $\Delta_0, \ldots, \Delta_{l-1}$ are certain differential operators of the form (2) but with $p-1$ in place of p, of orders at most $0, \ldots, l-1$ respectively. Multiplying the two determinants by rows, we obtain

$$GW = \det\left(\sum_{\rho=0}^{l-1} \Delta_\mu \frac{1}{\nu\,!}\left(\frac{\partial}{\partial x_p}\right)^{\nu} \phi_\rho(x_p)\,\psi_\rho(x_1, \ldots, x_{p-1})\right)$$

$$= \det\left(\Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial x_p}\right)^{\nu} R\right) \quad (\mu, \nu = 0, \ldots, l-1).$$

Thus $W(x_p)\,G(x_1, \ldots, x_{p-1}) = F(x_1, \ldots, x_p)$, say, is representable in the form (6). It is plain from (6) that F has integral coefficients, and since W and G are not identically zero, neither is F.

From the fact that

$$F(x_1, \ldots, x_p) = W(x_p)\,G(x_1, \ldots, x_{p-1}),$$

where F has integral coefficients and W, G have rational coefficients, it follows that there exists a rational number g such that the polynomials $U(x_1, \ldots, x_{p-1}) = gG(x_1, \ldots, x_{p-1})$ and $V(x_p) = g^{-1}W(x_p)$ have integral coefficients†.

Finally, since W is a determinant of order l whose elements are polynomials in x_p of degree r_p at most, it follows that W, and therefore V, is a polynomial in x_p of degree lr_p at most. Similarly G, and therefore U, is of degree at most lr_j in x_j for $j = 1, \ldots, p-1$.

We have now proved all that was asserted.

---

† See, for example, Perron, *Algebra* I (Berlin, 1927, 1931, 1951), Satz 88. The deduction does not depend on the separation of the variables between G and W.

LEMMA 3. *Let R satisfy the hypotheses of Lemma* 2 *and suppose that all the coefficients of R have absolute values not exceeding B. Then all the coefficients of $F(x_1, \ldots, x_p)$, defined in* (6), *have absolute values not exceeding*

$$\big((r_1+1)\ldots(r_p+1)\big)^l\, l!\, B^l\, 2^{(r_1+\ldots+r_p)l}.$$

Proof. In the definition (6) of F, we can regard R as a sum of $(r_1+1)\ldots(r_p+1)$ terms, each of the form

$$a_{s_1,\ldots,s_p}\, x_1^{s_1}\ldots x_p^{s_p},$$

where $|a_{s_1,\ldots,s_p}| \leqslant B$. The determinant on the right of (6) can be developed into a sum of $\big((r_1+1)\ldots(r_p+1)\big)^l$ determinants, the general element in one such determinant being of the form

$$a_{s_1,\ldots,s_p}\, \Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial x_p}\right)^\nu x_1^{s_1}\ldots x_p^{s_p},$$

where $s_1, \ldots, s_p$ depend on μ, or alternatively on ν, according as the original determinant is developed by rows or columns. Now

$$\Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial x_p}\right)^\nu x_1^{s_1}\ldots x_p^{s_p} = A x_1^{t_1}\ldots x_p^{t_p}$$

for some $t_1 \leqslant s_1, \ldots, t_p \leqslant s_p$, and the coefficient A, if not zero, is given by

$$A = \binom{s_1}{t_1}\ldots\binom{s_p}{t_p}.$$

Thus $$A \leqslant 2^{s_1+\ldots+s_p} \leqslant 2^{r_1+\ldots+r_p}.$$

Hence the coefficients of each of the $l!$ terms in the expansion of an individual determinant have absolute values not exceeding

$$(B2^{r_1+\ldots+r_p})^l,$$

and the result follows.

4. Let $P(x_1, \ldots, x_p)$ be any polynomial in p variables which does not vanish identically. Let $\alpha_1, \ldots, \alpha_p$ be any real numbers, and let $r_1, \ldots, r_p$ be any positive numbers. We define *the index θ of P at the point $(\alpha_1, \ldots, \alpha_p)$ relative to $r_1, \ldots, r_p$* as follows. Expand $P(\alpha_1+y_1, \ldots, \alpha_p+y_p)$ as a polynomial in $y_1, \ldots, y_p$, say

$$P(\alpha_1+y_1, \ldots, \alpha_p+y_p) = \sum_{j_1=0}^{\infty}\ldots\sum_{j_p=0}^{\infty} c(j_1, \ldots, j_p)\, y_1^{j_1}\ldots y_p^{j_p}.$$

Then $$\theta = \min\left(\frac{j_1}{r_1}+\ldots+\frac{j_p}{r_p}\right)$$

for all sets of non-negative integers $j_1, \ldots, j_p$ for which

$$c(j_1, \ldots, j_p) \neq 0.$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

The last condition can obviously be expressed equivalently as

$$\left(\frac{\partial}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial x_p}\right)^{j_p} P(\alpha_1, \ldots, \alpha_p) \neq 0.$$

We note that $\theta \geqslant 0$ always, and $\theta = 0$ if and only if $P(\alpha_1, \ldots, \alpha_p) \neq 0$. We note also that the index of the derived polynomial

$$\left(\frac{\partial}{\partial x_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial x_p}\right)^{k_p} P(x_1, \ldots, x_p)$$

at $(\alpha_1, \ldots, \alpha_p)$ relative to $r_1, \ldots, r_p$ is at least

$$\theta - \frac{k_1}{r_1} - \ldots - \frac{k_p}{r_p}$$

for any non-negative integers $k_1, \ldots, k_p$, provided that the derived polynomial is not identically zero. Some further immediate consequences of the definition are given in the following lemma.

LEMMA 4. *Let $P(x_1, \ldots, x_p)$ and $Q(x_1, \ldots, x_p)$ be polynomials, neither of which vanishes identically. Then, if all the indices are formed at the same point $(\alpha_1, \ldots, \alpha_p)$ relative to the same numbers $r_1, \ldots, r_p$, we have*

$$\text{index } (P+Q) \geqslant \min (\text{index } P, \text{ index } Q), \tag{8}$$

$$\text{index } PQ = \text{index } P + \text{index } Q. \tag{9}$$

(9) *remains true if P is a polynomial in $x_1, \ldots, x_{p-1}$ only and Q is a polynomial in x_p only, and the index of P is taken at $(\alpha_1, \ldots, \alpha_{p-1})$ relative to $r_1, \ldots, r_{p-1}$ and that of Q at α_p relative to r_p.*

5. We consider, for a particular set of positive integers $r_1, \ldots, r_m$ and a particular number $B \geqslant 1$, polynomials $R(x_1, \ldots, x_m)$ in m variables which satisfy the conditions:

(*a*) R has integral coefficients and is not identically zero;

(*b*) R is of degree at most r_j in x_j for $j = 1, \ldots, m$;

(*c*) the coefficients of R have absolute values not exceeding B.

We denote the aggregate of all such polynomials by

$$\mathfrak{R}_m = \mathfrak{R}_m(B;\ r_1, \ldots, r_m).$$

Let $q_1, \ldots, q_m$ denote positive integers and let $h_1, \ldots, h_m$ denote integers satisfying $(h_j, q_j) = 1$ for $j = 1, \ldots, m$. Let $\theta(R)$ denote the index of $R(x_1, \ldots, x_m)$ at the point $(h_1/q_1, \ldots, h_m/q_m)$ relative to $r_1, \ldots, r_m$. Our object in the present section is to obtain, under certain conditions, an estimate for $\theta(R)$ in terms of $B, q_1, \ldots, q_m, r_1, \ldots, r_m$. We therefore define

$$\Theta_m(B;\ q_1, \ldots, q_m;\ r_1, \ldots, r_m) = \text{upper bound of } \theta(R) \tag{10}$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

taken over all polynomials R in the set $\mathfrak{R}_m$ and over all integers $h_1, \ldots, h_m$ which are relatively prime to $q_1, \ldots, q_m$ respectively.

It is important to observe the double significance of $r_1, \ldots, r_m$ in the definition (10); these numbers occur both in the definition of the index $\theta(R)$ and in condition (*b*) above.

Our arguments are based on induction with respect to m, and in the course of the work we shall need to use the above definitions for various values of m and for various sets of values of $B, q_1, \ldots, q_m, r_1, \ldots, r_m$.

The case $m = 1$ is simple, and can be treated without imposing any new conditions.

LEMMA 5. *We have*

$$\Theta_1(B;\ q_1;\ r_1) \leqslant \frac{\log B}{r_1 \log q_1}. \tag{11}$$

Proof. By the definition of the index θ of R, the polynomial $R(x_1)$ is divisible by† $(x_1 - h_1/q_1)^{\theta r_1}$. It follows from Gauss's theorem on the factorization of polynomials with integral coefficients into polynomials with rational coefficients, and from the fact that $(h_1, q_1) = 1$, that

$$R(x_1) = (q_1 x_1 - h_1)^{\theta r_1} Q(x_1),$$

where $Q(x_1)$ is a polynomial with integral coefficients. Hence the coefficient of the highest term in $R(x_1)$ is an integral multiple of $q_1^{\theta r_1}$, so that

$$q_1^{\theta r_1} \leqslant B,$$

giving (11). [It may be noted in passing that we have not used the hypothesis that the degree of R is at most r_1; the double significance of $r_1, \ldots, r_m$ mentioned above becomes important only when $m > 1$.]

We now come to the inductive argument.

LEMMA 6. *Let $p \geqslant 2$ be a positive integer, let $r_1, \ldots, r_p$ be positive integers satisfying*

$$r_p > 10\delta^{-1}, \quad r_{j-1}/r_j > \delta^{-1} \text{ for } j = 2, \ldots, p, \tag{12}$$

where $0 < \delta < 1$, and let $q_1, \ldots, q_p$ be positive integers. Then

$$\Theta_p(B;\ q_1, \ldots, q_p;\ r_1, \ldots, r_p) \leqslant 2 \max_l (\Phi + \Phi^{1/2} + \delta^{1/2}), \tag{13}$$

where the maximum is taken over integers l satisfying

$$1 \leqslant l \leqslant r_p + 1, \tag{14}$$

† The exponent θr_1 is of course a non-negative integer, and can be supposed to be a positive integer.

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

and where

$$\Phi = \Theta_1(M;\, q_p;\, lr_p) + \Theta_{p-1}(M;\, q_1, \ldots, q_{p-1};\, lr_1, \ldots, lr_{p-1}) \qquad (15)$$

and $$M = (r_1+1)^{pl} l!\, B^l 2^{plr_1}. \qquad (16)$$

Proof. We have to show that if $R(x_1, \ldots, x_p)$ is any polynomial in the class $\mathfrak{R}_p(B;\, r_1, \ldots, r_p)$, and if $h_1, \ldots, h_p$ are integers relatively prime to $q_1, \ldots, q_p$ respectively, then the index θ of R at $(h_1/q_1, \ldots, h_p/q_p)$ relative to $r_1, \ldots, r_p$ does not exceed the right-hand side of (13).

The polynomial $R(x_1, \ldots, x_p)$ satisfies the hypotheses of Lemma 2, and therefore there exist an integer l satisfying (14) and a polynomial $F(x_1, \ldots, x_p)$ of the form (6) with the properties (i) and (ii) of Lemma 2. By Lemma 3 the coefficients of F have absolute values not exceeding

$$\big((r_1+1)\ldots(r_p+1)\big)^l l!\, B^l 2^{(r_1+\ldots+r_p)l} < M$$

by (16), since $r_1 > r_2 > \ldots > r_p$ by (12). Since

$$F = U(x_1, \ldots, x_{p-1})\, V(x_p),$$

and U, V have integral coefficients, it follows that the coefficients of U and V also have absolute values less than M.

The polynomial $U(x_1, \ldots, x_{p-1})$ has degree at most lr_j in x_j for $j = 1, \ldots, p-1$. It satisfies the conditions (*a*), (*b*), (*c*) above for the class of polynomials

$$\mathfrak{R}_{p-1}(M;\, lr_1, \ldots, lr_{p-1}).$$

Hence its index at $(h_1/q_1, \ldots, h_{p-1}/q_{p-1})$ *relative to* $lr_1, \ldots, lr_{p-1}$ does not exceed

$$\Theta_{p-1}(M;\, q_1, \ldots, q_{p-1};\, lr_1, \ldots, lr_{p-1}).$$

It follows from the definition of the index that the index of U at that point relative to $r_1, \ldots, r_{p-1}$ does not exceed

$$l\Theta_{p-1}(M;\, q_1, \ldots, q_{p-1};\, lr_1, \ldots, lr_{p-1}).$$

Similarly $V(x_p)$ belongs to the class $\mathfrak{R}_1(M;\, lr_p)$, and its index at h_p/q_p relative to r_p does not exceed

$$l\Theta_1(M;\, q_p;\, lr_p).$$

By the final clause of Lemma 4, the index of $F = UV$ at $(h_1/q_1, \ldots, h_p/q_p)$ relative to $r_1, \ldots, r_p$ is the sum of the indices of U and V, whence

$$\text{index } F \leqslant l\Phi, \qquad (17)$$

where Φ is defined in (15).

We now deduce from the determinantal representation of F in (6) a lower bound for the index of F in terms of the index θ of R. Consider first

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

any differential operator of the form

$$\Delta = \frac{1}{i_1!\dots i_{p-1}!}\left(\frac{\partial}{\partial x_1}\right)^{i_1}\dots\left(\frac{\partial}{\partial x_{p-1}}\right)^{i_{p-1}}$$

on $x_1, \dots, x_{p-1}$, of order $w = i_1+\dots+i_{p-1} \leqslant l-1$. If the polynomial

$$\Delta \frac{1}{\nu!}\left(\frac{\partial}{\partial x_p}\right)^{\nu} R(x_1, \dots, x_p)$$

does not vanish identically, its index at $(h_1/q_1, \dots, h_p/q_p)$ relative to $r_1, \dots, r_p$ is at least

$$\theta - \frac{i_1}{r_1} - \dots - \frac{i_{p-1}}{r_{p-1}} - \frac{\nu}{r_p} \geqslant \theta - \frac{w}{r_{p-1}} - \frac{\nu}{r_p}.$$

Now $w/r_{p-1} \leqslant (l-1)/r_{p-1} \leqslant r_p/r_{p-1} < \delta$, by (14) and (12). Hence, since the index is never negative, it must be at least

$$\max(0, \theta - \nu/r_p) - \delta.$$

If we expand the determinant on the right of (6), we obtain for F a sum of $l!$ terms, the typical term being of the form

$$\pm(\Delta_{\mu_0} R)\left(\Delta_{\mu_1}\frac{1}{1!}\frac{\partial}{\partial x_p}R\right)\dots\left(\Delta_{\mu_{l-1}}\frac{1}{(l-1)!}\left(\frac{\partial}{\partial x_p}\right)^{l-1}R\right),$$

where $\Delta_{\mu_0}, \dots, \Delta_{\mu_{l-1}}$ are differential operators on $x_1, \dots, x_{p-1}$ whose orders are at most $l-1$. By Lemma 4, the index of such a term (if it does not vanish identically) is at least

$$\sum_{\nu=0}^{l-1}\max(0, \theta-\nu/r_p) - l\delta.$$

Since F is a sum of such terms, it follows from Lemma 4 again that

$$\text{index } F \geqslant \sum_{\nu=0}^{l-1}\max(0, \theta-\nu/r_p) - l\delta.$$

We can suppose that $\theta r_p > 10$, for if not we have

$$\theta \leqslant 10 r_p^{-1} < \delta < 2\delta^{1/2},$$

and the desired inequality for θ then holds. If $\theta r_p < l$, we have

$$\sum_{\nu=0}^{l-1}\max(0, \theta-\nu/r_p) = r_p^{-1}\sum_{0\leqslant\nu\leqslant\theta r_p}(\theta r_p - \nu)$$
$$\geqslant \tfrac{1}{2} r_p^{-1}[\theta r_p]^2$$
$$> \tfrac{1}{3} r_p \theta^2.$$

If $\theta r_p \geqslant l$, we have

$$\sum_{\nu=0}^{l-1}\max(0, \theta-\nu/r_p) = \sum_{\nu=0}^{l-1}(\theta-\nu/r_p) \geqslant \tfrac{1}{2} l\theta.$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

Hence $$\text{index } F \geqslant \min(\tfrac{1}{2}l\theta,\ \tfrac{1}{3}r_p\theta^2) - l\delta. \tag{18}$$

Combining the inequalities (17) and (18), we obtain

$$\min(\tfrac{1}{2}l\theta,\ \tfrac{1}{3}r_p\theta^2) \leqslant l(\Phi+\delta).$$

Hence either $\theta \leqslant 2(\Phi+\delta)$, in which case θ satisfies the desired inequality, or

$$\tfrac{1}{3}r_p\theta^2 \leqslant l(\Phi+\delta) \leqslant (r_p+1)(\Phi+\delta).$$

Since $r_p+1 < \frac{4}{3}r_p$ by (12), the latter implies

$$\theta < 2(\Phi+\delta)^{1/2} \leqslant 2(\Phi^{1/2}+\delta^{1/2}).$$

This completes the proof of Lemma 6.

We next deduce an explicit result, in a form suitable for use later, by giving B a particular value and imposing further restrictions on the q's and r's.

LEMMA 7. *Let m be a positive integer and let δ satisfy*

$$0 < \delta < m^{-1}. \tag{19}$$

Let $r_1, \ldots, r_m$ be positive integers satisfying

$$r_m > 10\delta^{-1}, \quad r_{j-1}/r_j > \delta^{-1} \textit{ for } j = 2, \ldots, m. \tag{20}$$

Let $q_1, \ldots, q_m$ be positive integers satisfying

$$\log q_1 > \delta^{-1}m(2m+1), \tag{21}$$

$$r_j \log q_j \geqslant r_1 \log q_1 \textit{ for } j = 2, \ldots, m. \tag{22}$$

Then $$\Theta_m(q_1^{\delta r_1};\ q_1, \ldots, q_m;\ r_1, \ldots, r_m) < 10^m \delta^{(1/2)^m}. \tag{23}$$

Proof. We establish Lemma 7 by induction on m. If $m = 1$, Lemma 5 gives

$$\Theta_1(q_1^{\delta r_1};\ q_1;\ r_1) \leqslant \frac{\delta r_1 \log q_1}{r_1 \log q_1} = \delta \leqslant 10\delta^{1/2},$$

and we obtain (23) without using the hypotheses (20) and (21).

Now suppose that $p \geqslant 2$ is an integer, and that Lemma 7 is valid when $m = p-1$. We proceed to prove Lemma 7 when $m = p$. The hypotheses of Lemma 7 when $m = p$ are more stringent than those of Lemma 6, hence Lemma 6 is applicable. We now estimate first M in (16) and then Φ in (15).

We have

$$M = (r_1+1)^{pl}\, l!\, 2^{plr_1} q_1^{\delta l r_1} \leqslant \left((r_1+1)^p\, l\, 2^{pr_1} q_1^{\delta r_1}\right)^l.$$

Now $l \leqslant r_p+1 < r_1+1 \leqslant 2^{r_1}$. Hence

$$M < (2^{(2p+1)r_1} q_1^{\delta r_1})^l < (e^{(2p+1)r_1} q_1^{\delta r_1})^l.$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

By (21) with $m=p$, we have $2p+1<\delta p^{-1}\log q_1$, whence

$$M<q_1^{\delta_1 lr_1},$$

where
$$\delta_1=\delta(1+p^{-1}). \tag{24}$$

Thus
$$\Theta_1(M;\ q_p;\ lr_p)\leqslant\Theta_1(q_1^{\delta_1 lr_1};\ q_p;\ lr_p) \tag{25}$$

and

$$\Theta_{p-1}(M;\ q_1,\ \ldots,\ q_{p-1};\ lr_1,\ \ldots,\ lr_{p-1})$$
$$\leqslant\Theta_{p-1}(q_1^{\delta_1 lr_1};\ q_1,\ \ldots,\ q_{p-1};\ lr_1,\ \ldots,\ lr_{p-1}). \tag{26}$$

By Lemma 5, the right-hand side of (25) does not exceed

$$\frac{\log(q_1^{\delta_1 lr_1})}{lr_p\log q_p}\leqslant\frac{\delta_1 lr_1\log q_1}{lr_1\log q_1}=\delta_1,$$

in view of (22).

To estimate the right-hand side of (26) we use the inductive hypothesis of the present proof, namely that Lemma 7 holds when $m=p-1$. The conditions of Lemma 7 for $m=p-1$ are satisfied when we replace δ by δ_1 and $r_1,\ \ldots,\ r_{p-1}$ by $lr_1,\ \ldots,\ lr_{p-1}$; since $\delta_1>\delta$ this is immediate for all but (19). To verify the analogue of (19) we have to show that

$$\delta_1<(p-1)^{-1},$$

and this follows from (24) and the fact that $\delta<p^{-1}$ by (19) with $m=p$. It follows that

$$\Theta_{p-1}(q_1^{\delta_1 lr_1};\ q_1,\ \ldots,\ q_{p-1};\ lr_1,\ \ldots,\ lr_{p-1})<10^{p-1}\delta_1^{(1/2)^{p-1}}.$$

Since $\delta_1<2\delta$, the two results just proved imply that

$$\Phi<2\delta+2(10^{p-1}\delta^{(1/2)^{p-1}})<3(10^{p-1}\delta^{(1/2)^{p-1}}).$$

Now (13) gives

$$\Theta_p(q_1^{\delta r_1};\ q_1,\ \ldots,\ q_p;\ r_1,\ \ldots,\ r_p)<2\left(3(10^{p-1}\delta^{(1/2)^{p-1}})+3^{1/2}\,10^{(p-1)/2}\,\delta^{(1/2)^p}+\delta^{1/2}\right)$$
$$<2\left(\frac{3}{10}+\frac{3^{1/2}}{10^{3/2}}+\frac{1}{10^2}\right)10^p\,\delta^{(1/2)}$$
$$<10^p\,\delta^{(1/2)^p}.$$

Thus Lemma 7 holds when $m=p$, as asserted.

6. The next lemma is independent of any hypotheses concerning the positive integers $r_1,\ \ldots,\ r_m$.

LEMMA 8. *If $r_1,\ \ldots,\ r_m$ are any positive integers, and $\lambda>0$, then the number of sets of integers $j_1,\ \ldots,\ j_m$ which satisfy the inequalities*

$$0\leqslant j_1\leqslant r_1,\ \ldots,\ 0\leqslant j_m\leqslant r_m,\quad \frac{j_1}{r_1}+\ldots+\frac{j_m}{r_m}\leqslant\tfrac{1}{2}(m-\lambda)$$

does not exceed
$$2m^{1/2}\lambda^{-1}(r_1+1)\ldots(r_m+1).$$

Proof. The result holds when $m=1$, for the number of integers j_1 satisfying

$$0 \leqslant j_1 \leqslant r_1, \quad j_1 \leqslant \tfrac{1}{2}(1-\lambda) r_1$$

is at most r_1+1 and is 0 if $\lambda > 1$.

We suppose $m>1$ and prove the result by induction on m. The result is trivial if $\lambda \leqslant 2m^{1/2}$, so we can suppose $\lambda > 2m^{1/2}$. For a particular value of j_m, the conditions on $j_1, \ldots, j_{m-1}$ are of the same general nature as before but with $m-1$ in place of m and with λ replaced by λ', where

$$\tfrac{1}{2}(m-1-\lambda') = \tfrac{1}{2}(m-\lambda) - j_m/r_m,$$

that is,
$$\lambda' = \lambda - 1 + 2j_m/r_m.$$

We note that $\lambda' > 0$ for $0 \leqslant j_m \leqslant r_m$, since $\lambda > 2m^{1/2} > 1$. By the hypothesis of the induction, the number of solutions of the original inequalities in $j_1, \ldots, j_m$ does not exceed

$$\sum_{j_m=0}^{r_m} 2(m-1)^{1/2} (\lambda - 1 + 2j_m/r_m)^{-1} (r_1+1) \ldots (r_{m-1}+1).$$

Hence it suffices to prove that

$$\sum_{j=0}^{r} (\lambda - 1 + 2j/r)^{-1} < \lambda^{-1}(m-1)^{-1/2} m^{1/2}(r+1)$$

for any positive integers r and m, when $\lambda > 2m^{1/2}$.

If we suppose r even, and replace j by $\tfrac{1}{2}r+k$, the sum becomes

$$\sum_{k=-r/2}^{r/2} (\lambda + 2k/r)^{-1} = \lambda^{-1} + \sum_{k=1}^{r/2} 2\lambda(\lambda^2 - 4k^2/r^2)^{-1}$$

$$\leqslant \lambda^{-1} + \sum_{k=1}^{r/2} 2\lambda(\lambda^2-1)^{-1}$$

$$\leqslant (r+1)\lambda^{-1}(1-\lambda^{-2})^{-1}.$$

Now $1-\lambda^{-2} > 1-\tfrac{1}{4}m^{-1} > (1-m^{-1})^{1/2}$, whence the result. A similar but slightly simpler argument applies if r is odd†.

7. Let α be a real algebraic number, not rational, and suppose that the inequality (1) is satisfied by infinitely many pairs of integers h, q with $q>0$. We can suppose that α is an algebraic integer; for if not there is a rational integer M such that $M\alpha$ is an algebraic integer, and the inequality

$$\left| M\alpha - \frac{h'}{q} \right| < \frac{M}{q^{\kappa}}$$

is satisfied by infinitely many pairs of integers h', q. Hence $M\alpha$ has the

† The case of even r would in fact suffice for the application later, since we could choose $r_1, \ldots, r_m$ in §8 so as to be even.

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

same property as α provided that in (1) we replace κ by any smaller number.

If α is an algebraic integer, there is some polynomial

$$f(x) = x^n + a_1 x^{n-1} + \ldots + a_n, \tag{27}$$

with integral coefficients and highest coefficient 1, such that $f(\alpha) = 0$. Put

$$A = \max(1, |a_1|, \ldots, |a_n|). \tag{28}$$

In the remainder of the paper we shall be concerned with only one set of values of m, δ, q_1, h_1, ..., q_m, h_m, r_1, ..., r_m, which will be chosen later in the order just indicated. The choice will be so made as to satisfy the following conditions:

$$0 < \delta < m^{-1}, \tag{29}$$

$$10^m \delta^{(1/2)^m} + 2(1+3\delta) n m^{1/2} < \tfrac{1}{2}m, \tag{30}$$

$$r_m > 10\delta^{-1}, \quad r_{j-1}/r_j > \delta^{-1} \text{ for } j = 2, \ldots, m, \tag{31}$$

$$\delta^2 \log q_1 > 2m+1+2m\log(1+A)+2m\log(1+|\alpha|), \tag{32}$$

$$r_j \log q_j \geqslant r_1 \log q_1. \tag{33}$$

We note that these conditions imply those of Lemma 7, since (29) and (32) imply $\delta \log q_1 > m(2m+1)$.

Define λ, γ, η, B_1 by

$$\lambda = 4(1+3\delta) n m^{1/2}, \tag{34}$$

$$\gamma = \tfrac{1}{2}(m-\lambda), \tag{35}$$

$$\eta = 10^m \delta^{(1/2)^m}, \tag{36}$$

$$B_1 = [q_1^{\delta r_1}]. \tag{37}$$

We note that (30) is equivalent to

$$\eta < \gamma. \tag{38}$$

We note also that B_1 is necessarily large, since $r_1 > 10$ and $q_1^{\delta^2} > e^{2m+1} \geqslant e^3$. Thus, in particular, $q_1^{\frac{1}{2}\delta r_1} < B_1$.

We now come to the main lemma, which is the only lemma to which reference will be made in the final proof of the theorem.

LEMMA 9. *Suppose the conditions* (29)–(33) *are satisfied, and suppose that* $h_1, \ldots, h_m$ *are integers relatively prime to* $q_1, \ldots, q_m$ *respectively. Then there exists a polynomial* $Q(x_1, \ldots, x_m)$ *with integral coefficients, of degree at most* r_j *in* x_j *for* $j = 1, \ldots, m$, *such that*

(i) *the index of* Q *at the point* $(\alpha, \ldots, \alpha)$ *relative to* $r_1, \ldots, r_m$ *is at least* $\gamma - \eta$;

(ii) $Q(h_1/q_1, \ldots, h_m/q_m) \neq 0$;

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

(iii) *for all derivatives*

$$Q_{i_1, \ldots, i_m}(x_1, \ldots, x_m) = \frac{1}{i_1! \ldots i_m!} \left(\frac{\partial}{\partial x_1}\right)^{i_1} \ldots \left(\frac{\partial}{\partial x_m}\right)^{i_m} Q,$$

where $i_1, \ldots, i_m$ *are any non-negative integers, we have*

$$|Q_{i_1, \ldots, i_m}(\alpha, \ldots, \alpha)| < B_1^{1+3\delta}. \tag{39}$$

Proof. We consider all polynomials $W(x_1, \ldots, x_m)$ of the form

$$W(x_1, \ldots, x_m) = \sum_{s_1=0}^{r_1} \ldots \sum_{s_m=0}^{r_m} c(s_1, \ldots, s_m) x_1^{s_1} \ldots x_m^{s_m}, \tag{40}$$

where the coefficients $c(s_1, \ldots, s_m)$ assume independently all integral values satisfying

$$0 \leqslant c(s_1, \ldots, s_m) \leqslant B_1. \tag{41}$$

The number of such polynomials W is

$$N = (B_1+1)^r, \tag{42}$$

where for brevity we write

$$r = (r_1+1) \ldots (r_m+1). \tag{43}$$

For each such polynomial W we consider the derivatives

$$W_{j_1, \ldots, j_m}(x_1, \ldots, x_m) = \frac{1}{j_1! \ldots j_m!} \left(\frac{\partial}{\partial x_1}\right)^{j_1} \ldots \left(\frac{\partial}{\partial x_m}\right)^{j_m} W$$

for all integers $j_1, \ldots, j_m$ satisfying

$$0 \leqslant j_1 \leqslant r_1, \ldots, 0 \leqslant j_m \leqslant r_m, \quad \frac{j_1}{r_1} + \ldots + \frac{j_m}{r_m} \leqslant \gamma. \tag{44}$$

By Lemma 8 and (35), the number D of such derivatives satisfies

$$D \leqslant 2m^{1/2} \lambda^{-1} r, \tag{45}$$

where r is given by (43).

For each such derivative we form the polynomial

$$W_{j_1, \ldots, j_m}(x, \ldots, x)$$

in a single variable x, and divide this polynomial by $f(x)$, denoting the remainder by

$$T_{j_1, \ldots, j_m}(W; x).$$

This remainder is a polynomial in x with integral coefficients, of degree $n-1$ at most.

We proceed to obtain an estimate for the magnitude of the coefficients in any such remainder. The coefficients in each derived polynomial $W_{j_1, \ldots, j_m}(x_1, \ldots, x_m)$ have absolute values not exceeding

$$2^{r_1+\ldots+r_m} B_1 \leqslant 2^{mr_1} B_1 < B_1^{1+\delta},$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

since $mr_1 \log 2 < \frac{1}{2}\delta^2 r_1 \log q_1$ by (32). When $x_1, \ldots, x_m$ are all replaced by x, some of the terms in the polynomial may coalesce; since the total number of terms is at most r, the coefficients in $W_{j_1, \ldots, j_m}(x, \ldots, x)$ have absolute values less than $rB_1^{1+\delta}$. Now

$$r = (r_1+1) \ldots (r_m+1) \leqslant 2^{r_1+\ldots+r_m} \leqslant 2^{mr_1} < B_1^{\delta},$$

so that $rB_1^{1+\delta} < B_1^{1+2\delta}$. It remains to consider the operation of dividing this polynomial, say

$$w_s x^s + w_{s-1} x^{s-1} + \ldots + w_0,$$

by $f(x)$, given in (27). The first operation (supposing $s \geqslant n$) is to subtract $w_s x^{s-n} f(x)$; and this gives a new polynomial whose coefficients are either of the form $w_\nu - a_{s-\nu} w_s$ or of the form w_ν. Hence the coefficients of the new polynomial have absolute values less than $(1+A)\, B_1^{1+2\delta}$, with A as in (28). The same consideration applies to the subsequent operations in the division process, and leads to the conclusion that the coefficients in the remainders $T_{j_1, \ldots, j_m}(W;\, x)$ have absolute values less than

$$(1+A)^{s-n+1} B_1^{1+2\delta}.$$

Since $s \leqslant r_1 + \ldots + r_m \leqslant mr_1$, this is less than

$$(1+A)^{mr_1} B_1^{1+2\delta} < B_1^{1+3\delta}$$

by (32).

In view of this estimate for the coefficients in each remainder T, the number of distinct sets of D remainders that can arise is less than

$$(1+2B_1^{1+3\delta})^{nD}.$$

By (45) and the definition of λ in (34), we have

$$(1+3\delta)\, nD \leqslant 2(1+3\delta)\, nm^{1/2} \lambda^{-1} r = \tfrac{1}{2} r,$$

whence $$(1+2B_1^{1+3\delta})^{nD} < (2+2B_1)^{r/2} < (1+B_1)^r.$$

By reference to (42), we see that the number of distinct possible sets of remainders is less than the number of polynomials W under consideration. Hence there exist two distinct polynomials, say W' and W'', of the form (40) such that

$$W'_{j_1, \ldots, j_m}(x, \ldots, x) - W''_{j_1, \ldots, j_m}(x, \ldots, x)$$

is divisible by $f(x)$ for all $j_1, \ldots, j_m$ satisfying (44). Putting $W^* = W' - W''$, we deduce that all the corresponding derivatives

$$W^*_{j_1, \ldots, j_m}(x_1, \ldots, x_m)$$

are zero when $x_1 = \ldots = x_m = \alpha$. Hence the index of W^* at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least γ. Also the coefficients of W^* are integers, not all zero, of absolute values not exceeding B_1.

We now appeal to Lemma 7, the conditions of which are satisfied, as was noted earlier. The polynomial $W^*(x_1, \ldots, x_m)$ satisfies the conditions

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

(*a*), (*b*), (*c*) of §5 and so belongs to the class

$$\mathfrak{R}_m(q_1^{\delta r_1};\ r_1,\ \ldots,\ r_m).$$

By Lemma 7, its index at $(h_1/q_1, \ldots, h_m/q_m)$ relative to $r_1, \ldots, r_m$ is less than η, defined in (36). Hence W^* possesses some derivative

$$Q(x_1, \ldots, x_m) = \frac{1}{k_1! \ldots k_m!}\left(\frac{\partial}{\partial x_1}\right)^{k_1} \ldots \left(\frac{\partial}{\partial x_m}\right)^{k_m} W^*,$$

with
$$\frac{k_1}{r_1} + \ldots + \frac{k_m}{r_m} < \eta,$$

such that
$$Q(h_1/q_1, \ldots, h_m/q_m) \neq 0.$$

The index of Q at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least $\gamma - \eta$. Thus Q has the properties (i) and (ii) of the enunciation.

Since the coefficients of W^* have absolute values at most B_1, it follows that the coefficients of Q have absolute values at most

$$2^{r_1+\ldots+r_m} B_1 \leqslant 2^{mr_1} B_1 < B_1^{1+\delta}.$$

Hence the coefficients of any further derivative

$$Q_{i_1, \ldots, i_m}(x_1, \ldots, x_m)$$

have absolute values less than $2^{mr_1} B_1^{1+\delta} < B_1^{1+2\delta}$. It follows that

$$|Q_{i_1, \ldots, i_m}(\alpha, \ldots, \alpha)| < B_1^{1+2\delta}(1+|\alpha|)^{r_1+\ldots+r_m},$$

and this implies (iii) since

$$(1+|\alpha|)^{mr_1} < B_1^{\delta}$$

by (32). This completes the proof of Lemma 9.

8. *Completion of the proof.* We suppose that $\kappa > 2$ and that the inequality

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^\kappa} \tag{46}$$

has infinitely many solutions in integers h, q with $q > 0$. Since α is irrational there must be infinitely many solutions with $(h, q) = 1$. We shall deduce a contradiction.

We first choose m so large that $m > 4nm^{1/2}$ and

$$\frac{2m}{m - 4nm^{1/2}} < \kappa, \tag{47}$$

as is possible since $\kappa > 2$. For sufficiently small δ we have

$$m - 4(1+3\delta)nm^{1/2} - 2\eta > 0,$$

where η is given by (36) and is arbitrarily small with δ. This condition is the same as (30). We choose δ to satisfy this, and to satisfy (29), and further to satisfy

$$\frac{2m(1+4\delta)}{m-4(1+3\delta)\,nm^{1/2}-2\eta}<\kappa, \tag{48}$$

as is possible in view of (47). The inequality (48) is equivalent to

$$\frac{m(1+4\delta)}{\gamma-\eta}<\kappa, \tag{49}$$

by (34) and (35).

Having chosen m and δ, we now choose a solution h_1, q_1 of (46) with $(h_1, q_1)=1$ and with q_1 sufficiently large to satisfy (32). We then choose further solutions $h_2, q_2;\ \ldots;\ h_m, q_m$, with $(h_j, q_j)=1$ throughout, to satisfy

$$\frac{\log q_j}{\log q_{j-1}}>\frac{2}{\delta} \quad (j=2, \ldots, m). \tag{50}$$

We now take r_1 to be any integer satisfying

$$r_1>\frac{10\log q_m}{\delta\log q_1}, \tag{51}$$

and define $r_2, \ldots, r_m$ by

$$\frac{r_1\log q_1}{\log q_j}\leqslant r_j<1+\frac{r_1\log q_1}{\log q_j} \quad (j=2, \ldots, m). \tag{52}$$

Then (33) is satisfied. Also

$$\frac{r_j\log q_j}{r_1\log q_1}<1+\frac{\log q_j}{r_1\log q_1}\leqslant 1+\frac{\log q_m}{r_1\log q_1}<1+\tfrac{1}{10}\delta. \tag{53}$$

The conditions (31) are satisfied, since

$$r_m\geqslant\frac{r_1\log q_1}{\log q_m}>10\delta^{-1}$$

and

$$\frac{r_{j-1}}{r_j}>\frac{\log q_j}{\log q_{j-1}}(1+\tfrac{1}{10}\delta)^{-1}>\delta^{-1}$$

by (52), (53) and (50).

By Lemma 9 there exists a polynomial $Q(x_1, \ldots, x_m)$ with the properties stated there. The contradiction is reached by comparing two inequalities for $Q(h_1/q_1, \ldots, h_m/q_m)$, which is not 0 by (ii) of Lemma 9. Since Q has integral coefficients and is of degree at most r_j in x_j for $j=1, \ldots, m$, we have

$$|Q(h_1/q_1, \ldots, h_m/q_m)|\geqslant q_1^{-r_1}\ldots q_m^{-r_m}>q_1^{-mr_1(1+\delta)} \tag{54}$$

by (53). On the other hand, we have

$$Q(h_1/q_1, \ldots, h_m/q_m)=\sum_{i_1=0}^{r_1}\ldots\sum_{i_m=0}^{r_m}Q_{i_1,\ldots,i_m}(\alpha, \ldots, \alpha)(h_1/q_1-\alpha)^{i_1}\ldots(h_m/q_m-\alpha)^{i_m},$$

https://doi.org/10.1112/S0025579300000644 Published online by Cambridge University Press

and by (i) of Lemma 9 the terms with

$$\frac{i_1}{r_1}+\ldots+\frac{i_m}{r_m}<\gamma-\eta$$

all vanish. In every other term we have

$$\left|\left(\frac{h_1}{q_1}-\alpha\right)^{i_1}\ldots\left(\frac{h_m}{q_m}-\alpha\right)^{i_m}\right|<\frac{1}{(q_1^{i_1}\ldots q_m^{i_m})^{\kappa}}\leqslant q_1^{-r_1(\gamma-\eta)\kappa},$$

since $q_j\geqslant q_1^{r_1/r_j}$ by (52). Hence, using (iii) of Lemma 9, we have

$$|Q(h_1/q_1,\ \ldots,\ h_m/q_m)|<(r_1+1)\ldots(r_m+1)\,B_1^{1+3\delta}\,q_1^{-r_1(\gamma-\eta)\kappa}$$
$$<B_1^{1+4\delta}\,q_1^{-r_1(\gamma-\eta)\kappa}$$
$$<q_1^{(1+4\delta)\,\delta r_1-r_1(\gamma-\eta)\kappa}.$$

Comparing this with (54), we obtain

$$-mr_1(1+\delta)<(1+4\delta)\delta r_1-r_1(\gamma-\eta)\kappa,$$

or

$$\kappa<\frac{m(1+\delta)+\delta(1+4\delta)}{\gamma-\eta}<\frac{m(1+4\delta)}{\gamma-\eta},$$

contrary to (49). This completes the proof of the theorem.

University College,
London.

(*Received 26th January*, 1955. ; § 1 *revised 20th May*, 1955)