PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Contributions to the Theory of Diophantine Equations. I. On the Representation of Integers by Binary Forms

A. Baker

Phil. Trans. R. Soc. Lond. A 1968 **263**, 173-191
doi: 10.1098/rsta.1968.0010

CONTRIBUTIONS TO THE THEORY OF DIOPHANTINE EQUATIONS

I. ON THE REPRESENTATION OF INTEGERS BY BINARY FORMS

By A. BAKER
Trinity College, Cambridge

(*Communicated by H. Davenport, F.R.S.—Received* 4 *May* 1967)

CONTENTS

An effective algorithm is established for solving in integers x, y any Diophantine equation of the type $f(x, y) = m$, where f denotes an irreducible binary form with integer coefficients and degree at least 3. The magnitude, relative to m, of the bound furnished by the algorithm for the size of all the solutions of the equation is investigated, and, in consequence, there is obtained the first generally effective improvement on the well known result of Liouville (1844) concerning the accuracy with which algebraic numbers can be approximated by rationals.

1. Introduction

It was proved by Thue (1909) that the Diophantine equation

$$f(x,y) = m, \tag{1}$$

where f denotes an irreducible binary form with integer coefficients and degree at least 3, and m is any integer, possesses only a finite number of solutions in integers x, y. Thue discovered the theorem by way of his fundamental studies on rational approximations to algebraic numbers, which were later profoundly developed in the celebrated works of Siegel (1921) and Roth (1955)[†], and which formed the genesis of many other investigations. But Thue's theorem, like all subsequent developments, suffers from one basic limitation, that of its non-effectiveness. The proof depends on an assumption, made at the outset, that (1) possesses at least one solution in integers x, y with large absolute values, and the argument provides no way of deciding whether or not such a solution exists. A proof of Thue's theorem, in the case when $f(x, 1)$ has at least one complex zero, was given by Skolem (1935)[‡] by means of a $\mathfrak{p}$-adic argument very different from the original, but this also is of a non-effective character. Indeed it would seem that even for cubic polynomials f, no generally effective algorithm for the complete solution of (1) has hitherto been established, although

[†] See also Schneider (1936, 1957), Dyson (1947), Gelfond (1952).
[‡] Cf. Borevich & Shafarevich (1966, ch. 4).

a wide variety of techniques have been successfully employed to treat particular equations of this kind (cf. Skolem 1938).

The present paper is devoted to a new proof of Thue's theorem, which proceeds by an argument that is effective, and therefore provides an algorithm for the complete solution of (1) in integers x, y. Let $f(x, y)$ denote a homogeneous polynomial in x, y with degree $n \geqslant 3$ and with integer coefficients, irreducible over the rationals.† Suppose that

$$\kappa > n+1 \tag{2}$$

and let m be any positive‡ integer. The main result of this paper is as follows.

Theorem 1. *All solutions of* (1) *in integers x, y satisfy*

$$\max(|x|, |y|) < C\,\mathrm{e}^{(\log m)^{\kappa}},$$

where C is an effectively computable number depending only on n, κ and the coefficients of f.

The proof of the theorem depends essentially on the methods developed in two recent papers (Baker 1966, 1967) for the study of linear forms in the logarithms of algebraic numbers. It is necessary to modify the arguments of these papers to some extent, however, in order to apply them in the present context, and a self-contained exposition will therefore be given in which no knowledge of the earlier work is assumed. In fact a theorem on the logarithms of algebraic numbers of a slightly different nature to those obtained previously will first be proved (see § 2) and theorem 1 will then be established as a consequence of this result. The condition (2) above plays a similar rôle to the condition (1) of the first of the cited papers, and a relaxation in the latter condition, derived from an improvement in the basic techniques, would very likely lead to a corresponding (though not necessarily the same) relaxation in the former.||

From theorem 1 one obtains as an immediate deduction:

Theorem 2. *Suppose that α is an algebraic number with degree $n \geqslant 3$, and that $\kappa > n+1$. Then there is an effectively computable number $c = c(\alpha, \kappa) > 0$ such that*

$$\left|\alpha - \frac{p}{q}\right| > cq^{-n}\,\mathrm{e}^{(\log q)^{1/\kappa}} \tag{3}$$

for all integers p, q $(q > 0)$.

To verify this result it suffices to assume that α is real. If now (3) were not valid and if $f(x)$ denotes the minimal defining polynomial of α with relatively prime integer coefficients, then the mean-value theorem would give

$$q^n|f(p/q)| = q^n|f(p/q)-f(\alpha)| < c'q^n|\alpha - p/q| \leqslant cc'\,\mathrm{e}^{(\log q)^{1/\kappa}}$$

for some $c' = c'(\alpha) > 0$, and, since $y^n f(x/y)$ is a binary form with integer coefficients, this would clearly contradict theorem 1 if c is chosen sufficiently small.

Theorem 2 represents the first generally effective improvement on the well known result

† The terms 'form' and 'homogeneous polynomial' are used synonymously. f is said to be irreducible over the rationals if it cannot be expressed as the product of two binary forms with integers coefficients and degree less than n; by Gauss's lemma, f then also cannot be expressed as a product of binary forms with rational coefficients.

‡ The assumption that m is positive involves no loss of generality.

|| *Added in proof*, 9 *May* 1968: A third paper on linear forms in the logarithms of algebraic numbers has recently been published (*Mathematika* **14** (1967), 220–228) and it is shown here that the earlier requirement $\kappa > n+1$ can be relaxed to $\kappa > n$; the same improvement can be obtained in the present context.

of Liouville (1844). Liouville simply observed that $q^n|f(p/q)|$ above is a positive rational integer and so the number on the left of (3) exceeds cq^{-n} for some easily calculated $c = c(\alpha) > 0$. The much deeper works of Thue (1909), Siegel (1921) and Roth (1955) (see also the other papers cited earlier) show that cq^{-n} can be replaced by $cq^{-\kappa}$ for any $\kappa > 2$, but, as remarked earlier, the proofs are non-effective and do not enable this $c = c(\alpha, \kappa) > 0$ to be explicitly calculated. Some quantitative results in the direction of theorem 2 have previously been established for certain fractional powers of rationals (see Baker 1964 *a*, *b*), but here the arguments depend on particular properties of Gauss's hypergeometric function and the results would therefore seem to be of a rather special character. On the other hand, when applicable, the estimates obtained are stronger than those implied by theorem 2.

Theorem 1 may be regarded as the solution to a particular case of the tenth problem of Hilbert (1901). The question now arises as to how far the work here can be extended to give an effective algorithm applicable to other Diophantine equations in two unknowns. Siegel (1929) succeeded in generalizing Thue's original result and thereby established a simple necessary and sufficient condition for any equation of the form $F(x,y) = 0$, where F denotes a polynomial with integer coefficients, to have only a finite number of solutions in integers x, y. But Siegel's proof employs Weil's famous generalization of Mordell's finite basis theorem, and this again possesses a certain non-effective character. Thus it would seem that Siegel's work cannot lead directly to a quantitative extension of theorem 1.† Nevertheless, it is well known that many Diophantine equations in two unknowns can readily be reduced to a finite number of equations of the type (1), whence an effective algorithm is now available for their complete solution. These include, for example,

$$y^2 = x^3 + k,$$

where k is any integer, an equation which has long been a notable feature in the theory of numbers (cf. Mordell 1913, 1947, 1963). This will be the theme of Part II of the present paper.

2. On the logarithms of algebraic numbers

The purpose of this section is to give a precise formulation of the main result on linear forms in the logarithms of algebraic numbers which is fundamental to the proof of theorem 1. The following notation will be adopted. $\alpha_1, \dots, \alpha_n$ will denote $n \geqslant 2$ non-zero algebraic numbers. The maximum of the degrees of $\alpha_1, \dots, \alpha_n$ will be supposed not to exceed an integer d. $A_1, \dots, A_n$ will be used to denote the heights of $\alpha_1, \dots, \alpha_n$ respectively, that is the maximum of the absolute values of the relatively prime integer coefficients in their minimal defining polynomials. A will be written briefly for A_n, and A' will signify any number exceeding both $|\alpha_n|$ and $|\alpha_n|^{-1}$. $\log \alpha_1, \dots, \log \alpha_n$ will be understood to mean the principal values of the logarithms. It will be supposed that $\delta > 0$.

In § 7 it will be shown that the proof of theorem 1 can be reduced to the demonstration of the following result.

Theorem 3. *Suppose that $\kappa > n+1$ or $\kappa > n+2$ according as $\alpha_1, \dots, \alpha_n$ are or are not all real. Suppose further that $b_1, \dots, b_{n-1}$ are rational integers with absolute values at most H, such that*

$$0 < |\alpha_1^{b_1} \dots \alpha_{n-1}^{b_{n-1}} - \alpha_n| < e^{-\delta H}. \tag{4}$$

† A stronger version of (3) would also be required for the effective application of Siegel's argument.

Then
$$H < \max\{C, (\log A)^{\kappa}\}, \tag{5}$$
where C denotes an effectively computable number depending on n, d, δ, κ, $A_1, \ldots, A_{n-1}$ and A', but not on A.

We proceed now to prove that it suffices to establish a modified form of theorem 3 which relates to the logarithms of $\alpha_1, \ldots, \alpha_n$, in place of a product of powers. First, we note that for any complex number z, the inequality $|\mathrm{e}^z - 1| < \frac{1}{4}$ implies that
$$|z - \mathrm{i}k\pi| \leqslant 4|\mathrm{e}^z - 1| \tag{6}$$
for some rational integer k (i denoting, as usual, $(-1)^{\frac{1}{2}}$). For on writing $z = x + \mathrm{i}y$, where x and y are real, and putting $\eta = |\mathrm{e}^z - 1|$, we obtain
$$|\mathrm{e}^x \cos y - 1| \leqslant \eta, \quad |\mathrm{e}^x \sin y| \leqslant \eta.$$
The first inequality gives $|\mathrm{e}^x \cos y| \geqslant 1 - \eta$, and on combining with the second we get $|\tan y| \leqslant \eta/(1-\eta)$. Now there is a rational integer k such that $y' = |y - k\pi|$ satisfies $0 \leqslant y' \leqslant \frac{1}{2}\pi$ and, on noting that $x - \tan x$ decreases for $0 \leqslant x \leqslant \frac{1}{2}\pi$ and that $\eta < \frac{1}{4}$ we obtain
$$|y - k\pi| \leqslant \tan y' = |\tan y| \leqslant \eta/(1-\eta) \leqslant 2\eta.$$
This gives the required inequality (6), since obviously $|\mathrm{e}^x - 1| \leqslant \eta$ and so
$$|x| \leqslant \max(\log(1+\eta), \log(1-\eta)^{-1}) \leqslant \log(1+2\eta) \leqslant 2\eta.$$

It follows immediately from the last result that if (4) holds and if $\mathrm{e}^{\frac{1}{2}\delta H} > 4|\alpha_n|^{-1}$ then there is a rational integer b_0 such that
$$0 < |b_0 \log \alpha_0 + \ldots + b_{n-1} \log \alpha_{n-1} - \log \alpha_n| \leqslant 4|\alpha_n|^{-1} \mathrm{e}^{-\delta H} < \mathrm{e}^{-\frac{1}{2}\delta H},$$
where $\alpha_0 = -1$. Furthermore, if $\alpha_1, \ldots, \alpha_n$ are all real, then these inequalities hold with $b_0 = 0$, provided that $\alpha_1, \ldots, \alpha_n$ are replaced by $|\alpha_1|, \ldots, |\alpha_n|$ respectively. It is now clear that in order to prove theorem 3 it suffices to assume that n, $\alpha_1, \ldots, \alpha_n$, $A_1, \ldots, A_n$, d and δ are given as above, that κ satisfies (2), and thence to establish the assertion (5) under the new hypothesis that there exist rational integers $b_1, \ldots, b_{n-1}$ with absolute values at most H such that
$$0 < |b_1 \log \alpha_1 + \ldots + b_{n-1} \log \alpha_{n-1} - \log \alpha_n| < \mathrm{e}^{-\delta H}. \tag{7}$$

The inequality (7) is similar to the inequalities considered in Baker (1966, 1967) but no supposition has been made concerning the linear independence of $\log \alpha_1, \ldots, \log \alpha_n$ over the rationals, and this supposition played an essential rôle in the earlier work. It is therefore necessary to further modify the hypotheses of theorem 3, and the result which we shall finally prove is as follows.

THEOREM 4. *Suppose that $b_1, \ldots, b_n$ are rational integers with absolute values at most H^g (where H, g denote positive integers) such that*
$$0 < |b_1 \log \alpha_1 + \ldots + b_n \log \alpha_n| < \mathrm{e}^{-\delta H}. \tag{8}$$
Suppose further that there are no integers $b'_1, \ldots, b'_n$ with absolute values at most H for which
$$b'_1 \log \alpha_1 + \ldots + b'_n \log \alpha_n = 0, \tag{9}$$

other than $b'_1 = \ldots = b'_n = 0$. *Then* (5) *holds for some effectively computable number* C *depending on*† n, g, d, δ, κ, $A_1, \ldots, A_{n-1}$ *and* A', *but not on* A.

To show that theorem 3 follows from theorem 4 it is enough now to verify that from the existence of rational integers $b_1, \ldots, b_{n-1}$, with absolute values at most H, such that (7) holds, one can deduce the existence of an integer k between 1 and n inclusive, with the following properties: (i) There are rational integers $b_1, \ldots, b_n$, having absolute values at most $(2H)^{n-k+1}$, such that (8) is valid with $e^{-\delta H}$ on the right replaced by $H^{n-k}e^{-\delta H}$, (ii) at least $n-k$ of the integers $b_1, \ldots, b_n$ are 0, (iii) the only integers $b'_1, \ldots, b'_n$ with absolute values at most H such that (9) holds and such that also $b'_j = 0$ whenever $b_j = 0$, are given by $b'_1 = \ldots = b'_n = 0$. For if such an integer k exists then clearly the hypotheses of theorem 4 will be satisfied with n now given by the number of non-zero b_j, with g defined as twice the original n, and with δ replaced by $\frac{1}{2}\delta$, assuming, as we may, that H exceeds a sufficiently large number C as above. Note here that theorem 4 is applied to an arbitrary subset of the original $\alpha_1, \ldots, \alpha_n$ and the necessity to distinguish between $A = A_n$ and $A_1, \ldots, A_{n-1}$ does not always arise. To prove the existence of an integer k with the above properties we observe first that since, by virtue of (7), (i) and (ii) hold for $k = n$, it can be assumed that (iii) does not hold for $k = n$. Then there is a least positive integer k for which (i) and (ii) hold but (iii) does not, and obviously (ii) implies that $k > 1$ (assuming again that $H > C$). Let $b'_1, \ldots, b'_n$ denote a set of integers with the properties specified by (iii), other than $b'_1 = \ldots = b'_n = 0$. Suppose in fact that $b'_l \neq 0$ and put

$$b''_j = b_j b'_l - b'_j b_l \quad (1 \leqslant j \leqslant n),$$

where $b_1, \ldots, b_n$ denote integers satisfying (i) and (ii). Clearly (8) and (9) imply that

$$0 < |b''_1 \log \alpha_1 + \ldots + b''_n \log \alpha_n| < |b'_l|\, H^{n-k} e^{-\delta H} \leqslant H^{n-k+1} e^{-\delta H}.$$

Further we have $b''_j = 0$ whenever $b_j = 0$, and since also $b''_l = 0$, $b_l \neq 0$ we see that at least $n-k+1$ of the integers $b''_1, \ldots, b''_n$ are 0. Moreover, the b''_j have absolute values at most $(2H)^{n-k+2}$. Thus (i) and (ii) hold with k replaced by $k-1$. But, by the minimal choice of k, also (iii) must then hold with k replaced by $k-1$, and so $k-1$ has all the required properties.

3. Preliminaries to the proof of theorem 4

The next two sections will be devoted to the proof of theorem 4. In this section we introduce the notation that will be needed subsequently, and we record a few preliminary observations.

First, we note that if α is an algebraic number with degree d and height H then $|\alpha| \leqslant dH$. For if α satisfies an equation of the form

$$a_0 \alpha^d + a_1 \alpha^{d-1} + \ldots + a_d = 0,$$

† For the deduction of theorem 3 we shall require the observation that n here can be replaced by any integer n' satisfying $n' \geqslant n$ and $\kappa > n'+1$. In fact it will be seen that C can be expressed as a continuous function in the variables n, g, d, δ^{-1}, κ, $A_1, \ldots, A_{n-1}$, A' and $(\kappa-n-1)^{-1}$, monotonic increasing with respect to each. Further, we have $(\kappa-n-1)^{-1} \leqslant 1$ if there exists an integer n', other than n, with the above properties. Note also that a feature in the previous work on the logarithms of algebraic numbers which would have introduced a discontinuity into C has been avoided here (see the footnote on p. 213 of Baker 1966).

where the a_j denote integers with absolute values at most H and $a_0 \geqslant 1$, then either $|\alpha| < 1$ or

$$|\alpha| \leqslant |a_0 \alpha| = |a_1 + a_2 \alpha^{-1} + \ldots + a_d \alpha^{-d+1}| \leqslant dH.$$

Secondly, we observe that for each non-negative integer j we have

$$(a_0 \alpha)^j = a_0^{(j)} + a_1^{(j)} \alpha + \ldots + a_{d-1}^{(j)} \alpha^{d-1},$$

where the $a_m^{(j)}$ denote integers with absolute values at most $(2H)^j$; this follows easily from the recurrence relations

$$a_m^{(j)} = a_0 a_{m-1}^{(j-1)} - a_{d-m} a_{d-1}^{(j-1)} \quad (0 \leqslant m < d, j \geqslant d),$$

where $a_{-1}^{(j-1)} = 0$.

The notation introduced at the beginning of § 2 will be assumed without change, and it will further be supposed that g is a positive integer and that κ satisfies (2). $C, c_1, c_2, \ldots$ will denote numbers, greater than 1, which can be specified explicitly in terms of n, g, d, δ, κ, $A_1, \ldots, A_{n-1}$ and A' only. The number C, which will finally represent the constant occurring in the enunciation of theorem 4, will be supposed sufficiently large throughout. We assume now that the hypotheses of theorem 4 hold but that the conclusion is not valid, and we shall ultimately deduce a contradiction. Thus we assume that there exist rational integers $b_1, \ldots, b_n$ with absolute values at most H^g such that (8) holds, that the only integers $b'_1, \ldots, b'_n$ with absolute values at most H such that (9) holds are given by $b'_1 = \ldots = b'_n = 0$, and that

$$H \geqslant \max\{C, (\log A)^\kappa\}. \tag{10}$$

Further we assume that $b_n \neq 0$; this involves no loss of generality for, by (8), one at least of $b_1, \ldots, b_n$ is not 0, and, clearly, if $b_n = 0$ then, after modifying the notation, it would suffice to prove a less precise result in which C could possibly depend on A. For brevity we write

$$\beta_j = -b_j / b_n \quad (1 \leqslant j < n),$$

and we note that $\beta_1, \ldots, \beta_{n-1}$ satisfy

$$0 < |\beta_1 \log \alpha_1 + \ldots + \beta_{n-1} \log \alpha_{n-1} - \log \alpha_n| < \mathrm{e}^{-\delta H}. \tag{11}$$

Also since, for any complex number z,

$$|\mathrm{e}^z - 1| \leqslant |z|\, \mathrm{e}^{|z|}, \tag{12}$$

we obtain from (11)†

$$|\alpha_1^{\beta_1} \ldots \alpha_{n-1}^{\beta_{n-1}} - \alpha_n| < |\alpha_n|\, \mathrm{e}^{-\delta H + 1}. \tag{13}$$

We now define

$$\zeta = 2/(\kappa + n + 1), \quad \epsilon = \{1 - 1/(\zeta\kappa)\}/(2n),$$

and we observe that, by (2), we have

$$1/\kappa < \zeta < 1/(n+1) \quad \text{and} \quad 0 < \epsilon < 1/(2n).$$

We write

$$h = 8g[\log H], \quad k = [H^\zeta],$$

where, as usual, $[x]$ denotes the integral part of x. Further we define

$$L = L_1 = \ldots = L_{n-1} = [k^{1-\epsilon}], \quad L_n = [k^{n\epsilon}],$$

and we put $D = d^n$.

† It will be understood that $z^w = \mathrm{e}^{w \log z}$, where $\log z$ denotes the principal value of the logarithm.

Finally, for any integral function $f(z_1, \dots, z_{n-1})$ of the complex variables $z_1, \dots, z_{n-1}$ and any non-negative integers $m_1, \dots, m_{n-1}$ we write

$$f_{m_1, \dots, m_{n-1}}(z_1, \dots, z_{n-1}) = \frac{\partial^{m_1+\dots+m_{n-1}}}{\partial z_1^{m_1} \dots \partial z_{n-1}^{m_{n-1}}} f(z_1, \dots, z_{n-1}).$$

4. Lemmas

This section establishes six lemmas preliminary to the proof of theorem 4. As remarked earlier, the arguments given here are similar to those employed in Baker (1966, 1967), and indeed some parts of the discussion have been adopted without change. Nevertheless, proofs to all the lemmas have been supplied in detail.

Lemma 1. *Let M, N denote integers with $N > M > 0$ and let u_{ij} $(1 \leqslant i \leqslant M,\ 1 \leqslant j \leqslant N)$ denote integers with absolute values at most U. Then there exist integers $x_1, \dots, x_N$, not all 0, with absolute values at most $(NU)^{M/(N-M)}$ such that*

$$\sum_{j=1}^{N} u_{ij} x_j = 0 \quad (1 \leqslant i \leqslant M). \tag{14}$$

Proof. We put $B = [(NU)^{M/(N-M)}]$ and we note that there are $(B+1)^N$ different sets of integers $x_1, \dots, x_N$ with $0 \leqslant x_j \leqslant B$ $(1 \leqslant j \leqslant N)$. For each such set we have

$$-V_i B \leqslant y_i \leqslant W_i B \quad (1 \leqslant i \leqslant M),$$

where y_i denotes the left-hand side of (14), and $-V_i$, W_i denote the sum of the negative and positive u_{ij} $(1 \leqslant j \leqslant N)$ respectively. Since $V_i + W_i \leqslant NU$, there are at most $(NUB+1)^M$ different sets $y_1, \dots, y_M$. Now

$$(B+1)^{N-M} > (NU)^M$$

and so $(B+1)^N > (NUB+1)^M$. Hence there are two distinct sets $x_1, \dots, x_N$ which correspond to the same set $y_1, \dots, y_M$, and their difference gives the required solution of (14).

Lemma 2. *There are integers $p(\lambda_1, \dots, \lambda_n)$, not all 0, with absolute values at most e^{hk}, such that the function*

$$\Phi(z_1, \dots, z_{n-1}) = \sum_{\lambda_1=0}^{L_1} \dots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \dots, \lambda_n)\, \alpha_1^{\gamma_1 z_1} \dots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}},$$

where $\gamma_r = \lambda_r + \lambda_n \beta_r$ $(1 \leqslant r < n)$, satisfies

$$|\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)| < \mathrm{e}^{-\frac{1}{2}\delta H} \tag{15}$$

for all integers l with $1 \leqslant l \leqslant h$ and all non-negative integers $m_1, \dots, m_{n-1}$ with $m_1 + \dots + m_{n-1} \leqslant k$.

Proof. We shall determine the $p(\lambda_1, \dots, \lambda_n)$ such that

$$\sum_{\lambda_1=0}^{L_1} \dots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \dots, \lambda_n)\, \alpha_1^{\lambda_1 l} \dots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \dots \gamma_{n-1}^{m_{n-1}} = 0 \tag{16}$$

for the above ranges of l and $m_1, \dots, m_{n-1}$, and we shall verify subsequently that (16) implies (15). Let $a_1, \dots, a_n$ denote the leading coefficients (supposed positive) in the minimal defining polynomials of $\alpha_1, \dots, \alpha_n$ respectively. Then for any non-negative integer j we have

$$(a_r \alpha_r)^j = \sum_{s=0}^{d-1} a_{rs}^{(j)} \alpha_r^s, \tag{17}$$

where the $a_{rs}^{(j)}$ denote rational integers with absolute values at most c_1^j or $(2A)^j$ according as

$r < n$ or $r = n$ (see § 3). Thus multiplying (16) by

$$a_1^{L_1 l} \dots a_n^{L_n l}\, b_n^{m_1+\dots+m_{n-1}},$$

and substituting from (17) for the powers of $a_r \alpha_r$, we obtain the equation

$$\sum_{s_1=0}^{d-1} \dots \sum_{s_n=0}^{d-1} V(s)\, \alpha_1^{s_1} \dots \alpha_n^{s_n} = 0,$$

where
$$V(s) = \sum_{\lambda_1=0}^{L_1} \dots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \dots, \lambda_n)\, v(\lambda, s)$$

and
$$v(\lambda, s) = a_n^{(L_n-\lambda_n)l}\, a_{n,s_n}^{(\lambda_n l)} \prod_{r=1}^{n-1} \{a_r^{(L-\lambda_r)l}\, a_{r,s_r}^{(\lambda_r l)}\, (b_n \lambda_r + b_r \lambda_n)^{m_r}\}.$$

Hence (16) will be satisfied if the D equations $V(s) = 0$ hold. Now these represent linear equations in the $p(\lambda_1, \dots, \lambda_n)$ with integer coefficients. Further, since

$$l \leqslant h, \quad m_1 + \dots + m_{n-1} \leqslant k, \quad L_n < L$$

and, by hypothesis, the integers b_r have absolute values at most H^g, we deduce easily that the coefficient $v(\lambda, s)$ of $p(\lambda_1, \dots, \lambda_n)$ in the linear form $V(s)$ has absolute value at most

$$U = c_2^{Lh}\, (2A)^{L_n h}\, (2LH^g)^k.$$

Now there are at most $(k+1)^{n-1} h$ distinct sets of integers $l, m_1, \dots, m_{n-1}$, and hence there are $M \leqslant D(k+1)^{n-1} h$ equations $V(s) = 0$ corresponding to them. Further, there are $N = (L_1+1) \dots (L_n+1)$ unknowns $p(\lambda_1, \dots, \lambda_n)$ and

$$N > k^{(n-1)(1-\epsilon)+n\epsilon} = k^{n-1+\epsilon} > 2M,$$

since, clearly, k exceeds any fixed power of h if H is sufficiently large. It follows from Lemma 1 that the system of equations $V(s) = 0$ can be solved non-trivially, and indeed the integers $p(\lambda_1, \dots, \lambda_n)$ can be chosen to have absolute values at most NU. Now by (10) and the definitions introduced at the end of § 3 we obtain

$$L_n \log A \leqslant k^{n\epsilon} H^{1/\kappa} \leqslant k^{n\epsilon} (2k)^{1/(\zeta\kappa)} \leqslant 2k^{1-n\epsilon}, \tag{18}$$

and since also
$$2LH^g \leqslant 2kH^g \leqslant H^{g+1} \leqslant \mathrm{e}^{\frac{1}{2}h}, \tag{19}$$

it follows easily that
$$NU \leqslant k^n (2c_2)^{Lh}\, \mathrm{e}^{hL_n \log A + \frac{1}{2}hk} \leqslant \mathrm{e}^{hk},$$

as required.

It remains only to verify that (16) implies (15). Now it is clear that the left-hand side of (16) is obtained from $\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)$, apart from a factor

$$P = (\log \alpha_1)^{m_1} \dots (\log \alpha_{n-1})^{m_{n-1}} \tag{20}$$

in the latter, by substituting α_n for $\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}}$. From (13) we deduce that

$$|(\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}})^{\lambda_n l} - \alpha_n^{\lambda_n l}| \leqslant \lambda_n l(|\alpha_n| + 1)^{\lambda_n l}\, |\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}} - \alpha_n| \leqslant c_3^{L_n l}\, \mathrm{e}^{-\delta H} \leqslant c_3^{hk}\, \mathrm{e}^{-\delta H}.$$

Further we have
$$|P\alpha_1^{\lambda_1 l} \dots \alpha_{n-1}^{\lambda_{n-1} l}\, \gamma_1^{m_1} \dots \gamma_{n-1}^{m_{n-1}}| \leqslant c_4^{k+Ll} (2LH^g)^k \leqslant \mathrm{e}^{hk}.$$

Thus we obtain
$$|\Phi_{m_1, \ldots, m_{n-1}}(l, \ldots, l)| \leqslant (L_1+1)\ldots(L_n+1)\, \mathrm{e}^{2hk}\, c_3^{hk}\, \mathrm{e}^{-\delta H}$$
and (15) follows since $L_n < L < k$ and $hk < H^{\frac{1}{2}}$. This completes the proof of the lemma.

Lemma 3. *For any non-negative integers $m_1, \ldots, m_{n-1}$ with*
$$m_1+\ldots+m_{n-1} \leqslant k$$
and any complex number z we have
$$|\Phi_{m_1, \ldots, m_{n-1}}(z, \ldots, z)| < \mathrm{e}^{2hk}\, c_5^{L|z|}. \tag{21}$$
Further, for any integer l with $1 \leqslant l \leqslant hk^{(1/\zeta)+\frac{1}{2}\epsilon-1}$, either (15) *holds or*
$$|\Phi_{m_1, \ldots, m_{n-1}}(l, \ldots, l)| > (\mathrm{e}^{2hk}\, c_6^{Ll})^{-D}. \tag{22}$$

Proof. We have
$$\Phi_{m_1, \ldots, m_{n-1}}(z, \ldots, z) = P \sum_{\lambda_1=0}^{L_1} \ldots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \ldots, \lambda_n)\, q(\lambda, z),$$
where P is given by (20) and
$$q(\lambda, z) = \alpha_1^{\gamma_1 z} \ldots \alpha_{n-1}^{\gamma_{n-1} z}\, \gamma_1^{m_1} \ldots \gamma_{n-1}^{m_{n-1}}.$$
Now (11) implies that
$$|\alpha_1^{\beta_1 z} \ldots \alpha_{n-1}^{\beta_{n-1} z}| \leqslant \mathrm{e}^{(|\log \alpha_n|+1)|z|},$$
and since
$$|\log \alpha_n| \leqslant \max(|\alpha_n|, |\alpha_n|^{-1}) + 2\pi \leqslant A' + 2\pi,$$
and also, by (19), $|\gamma_r| \leqslant 2LH^g \leqslant \mathrm{e}^{\frac{1}{2}h}$, we obtain
$$|Pq(\lambda, z)| \leqslant \prod_{r=1}^{n-1} \{c_7^{L|z|}\, (c_8\, \mathrm{e}^{\frac{1}{2}h})^{m_r}\} \leqslant c_5^{L|z|}\, c_8^{k}\, \mathrm{e}^{\frac{1}{2}hk}.$$

Then (21) follows on noting that there are at most k^n terms in the above multiple sum and that the $p(\lambda_1, \ldots, \lambda_n)$ have absolute values at most e^{hk}.

To prove the second assertion we begin by defining
$$Q = P' \sum_{\lambda_1=0}^{L_1} \ldots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \ldots, \lambda_n)\, q'(\lambda, l),$$
where
$$P' = a_1^{L_1 l} \ldots a_n^{L_n l}\, b_n^{m_1+\ldots+m_{n-1}}, \quad q'(\lambda, l) = \alpha_1^{\lambda_1 l} \ldots \alpha_n^{\lambda_n l}\, \gamma_1^{m_1} \ldots \gamma_{n-1}^{m_{n-1}}$$
and the a_r are given as in the proof of lemma 2. Then it is clear that Q represents an algebraic integer with degree at most D. Further, on recalling the observation recorded at the beginning of § 3, applying (18), and noting also that, by (19),
$$|b_n^{m_1+\ldots+m_{n-1}}\, \gamma_1^{m_1} \ldots \gamma_{n-1}^{m_{n-1}}| \leqslant \mathrm{e}^{\frac{1}{2}hk},$$
we see that any conjugate of Q, obtained by substituting arbitrary conjugates for α_r, has absolute value at most
$$(L_1+1)\ldots(L_n+1)\, c_9^{Ll}\, (dA)^{2L_n l}\, \mathrm{e}^{\frac{3}{2}hk} \leqslant c_{10}^{Ll} \mathrm{e}^{2hk}.$$
Thus if $Q \neq 0$ we have $|\mathrm{norm}\, Q| \geqslant 1$ and hence
$$|Q| \geqslant (\mathrm{e}^{2hk}\, c_{10}^{Ll})^{-D+1}.$$
Now from (13) we obtain (cf. the end of the proof of lemma 2)
$$|q(\lambda, l) - q'(\lambda, l)| \leqslant c_{11}^{Ll}\, \mathrm{e}^{hk-\delta H},$$

and, by virtue of the inequalities $L \leqslant k^{1-\epsilon}$, $k \leqslant H^{\zeta}$ and the supposition $l \leqslant hk^{(1/\zeta)+\frac{1}{2}\epsilon-1}$, the number on the right is at most $\mathrm{e}^{-\frac{3}{4}\delta H}$. Hence we deduce that

$$|P^{-1}\Phi_{m_1, \ldots, m_{n-1}}(l, \ldots, l) - P'^{-1}Q| \leqslant (L_1+1)\ldots(L_n+1)\,\mathrm{e}^{hk-\frac{3}{4}\delta H} \leqslant \mathrm{e}^{-\frac{5}{8}\delta H}.$$

But since, again by (18) and (19),

$$|P'| \leqslant c_{12}^{Ll}\,A^{L_n l}H^{gk} \leqslant c_{13}^{Ll}\,\mathrm{e}^{hk},$$

and since also

$$c_{14}^{-k} < |P| < c_{15}^{k},$$

we see that either $Q = 0$ or

$$|PP'^{-1}Q| > (\mathrm{e}^{2hk}\,c_{16}^{Ll})^{-D} > 2\,\mathrm{e}^{-\frac{1}{2}\delta H};$$

further we have $P\,\mathrm{e}^{-\frac{5}{8}\delta H} < \mathrm{e}^{-\frac{1}{2}\delta H}$. This immediately establishes the second part of the lemma, the asserted alternatives corresponding to the cases $Q = 0$ or $Q \neq 0$.

Lemma 4. *Let J be any integer satisfying $0 \leqslant J < \tau$, where*

$$\tau = \epsilon^{-1}(n-1+\zeta^{-1})+1. \tag{23}$$

Then (15) *holds for all integers l with $1 \leqslant l \leqslant hk^{\frac{1}{2}\epsilon J}$ and each set of non-negative integers $m_1, \ldots, m_{n-1}$ with $m_1+\ldots+m_{n-1} \leqslant k/2^J$.*

Proof. The lemma is true for $J = 0$ by lemma 2. We suppose that K is an integer satisfying $0 \leqslant K \leqslant \tau-1$ and we assume that the lemma is true for $J = 0, 1, \ldots, K$. We proceed to prove the validity of the lemma for $J = K+1$.

We begin by defining

$$R_J = [hk^{\frac{1}{2}\epsilon J}], \quad S_J = [k/2^J] \quad (J = 0, 1, \ldots).$$

It suffices then to prove that for any integer l with $R_K < l \leqslant R_{K+1}$ and any set of non-negative integers $m_1, \ldots, m_{n-1}$ with $m_1+\ldots+m_{n-1} \leqslant S_{K+1}$ we have

$$|f(l)| < \mathrm{e}^{-\frac{1}{2}\delta H}, \tag{24}$$

where

$$f(z) = \Phi_{m_1, \ldots, m_{n-1}}(z, \ldots, z).$$

By our inductive hypothesis we see that for each integer r with $1 \leqslant r \leqslant R_K$ and each integer m satisfying $0 \leqslant m \leqslant S_{K+1}$ we have

$$|f_m(r)| < n^k\,\mathrm{e}^{-\frac{1}{2}\delta H}; \tag{25}$$

for $f_m(r)$ is given by

$$\left(\frac{\partial}{\partial z_1}+\ldots+\frac{\partial}{\partial z_{n-1}}\right)^m \Phi_{m_1, \ldots, m_{n-1}}(z_1, \ldots, z_{n-1})$$

evaluated at the point $z_1 = \ldots = z_{n-1} = r$, that is by

$$\sum_{j_1=0}^{m} \cdots \sum_{\substack{j_{n-1}=0 \\ j_1+\ldots+j_{n-1}=m}}^{m} m!\,(j_1!\ldots j_{n-1}!)^{-1}\,\Phi_{m_1+j_1, \ldots, m_{n-1}+j_{n-1}}(r, \ldots, r),$$

and the absolute values of the derivatives here are at most $\mathrm{e}^{-\frac{1}{2}\delta H}$ since

$$m_1+\ldots+m_{n-1}+j_1+\ldots+j_{n-1} \leqslant k/2^K.$$

We write, for brevity,

$$F(z) = \{(z-1)\ldots(z-R_K)\}^{S_{K+1}+1}$$

and we denote by Γ the circle in the complex plane, described in the positive sense, with centre the origin and with radius $R_{K+1}\log k$. Then by Cauchy's residue theorem we have

$$\frac{1}{2\pi \mathrm{i}}\int_{\Gamma}\frac{f(z)}{(z-l)\,F(z)}\,\mathrm{d}z=\frac{f(l)}{F(l)}+\frac{1}{2\pi \mathrm{i}}\sum_{r=1}^{R_K}\sum_{m=0}^{S_{K+1}}\frac{f_m(r)}{m!}\int_{\Gamma_r}\frac{(z-r)^m\,\mathrm{d}z}{(z-l)\,F(z)}, \tag{26}$$

where Γ_r denotes the circle in the complex plane, described in the positive sense, with centre r and radius $\frac{1}{2}$; for the residue of the pole of the integrand on the left at $z=r$ is given by

$$\frac{1}{S_{K+1}!}\frac{\mathrm{d}^{S_{K+1}}}{\mathrm{d}z^{S_{K+1}}}\left\{\frac{(z-r)^{S_{K+1}+1}f(z)}{(z-l)\,F(z)}\right\}$$

evaluated at $z=r$, and the integral over Γ_r on the right is given by

$$\frac{2\pi \mathrm{i}}{(S_{K+1}-m)!}\frac{\mathrm{d}^{S_{K+1}-m}}{\mathrm{d}z^{S_{K+1}-m}}\left\{\frac{(z-r)^{S_{K+1}+1}}{(z-l)\,F(z)}\right\}$$

again evaluated at $z=r$, and (26) now follows by Leibnitz's theorem. Since, for z on Γ_r,

$$|(z-r)^m/F(z)|<8^{S_{K+1}+1},$$

we deduce from (25) and the inequalities

$$R_K(S_{K+1}+1)\leqslant hk^{\frac{1}{2}\epsilon(\tau-1)+1}\leqslant hH^{\frac{1}{2}\{(n+1)\zeta+1\}}<H, \tag{27}$$

that the absolute value of the double sum on the right of (26) is at most

$$R_K(S_{K+1}+1)\,8^{S_{K+1}+2}n^k\,\mathrm{e}^{-\frac{1}{2}\delta H}\leqslant H(8n)^k\,\mathrm{e}^{-\frac{1}{2}\delta H}<\mathrm{e}^{-\frac{1}{4}\delta H}.$$

Further it is clear that

$$|F(l\leqslant)|\,l^{R_K(S_{K+1}+1)}\leqslant R_{K+1}^{R_K(S_{K+1}+1)}. \tag{28}$$

Also since $R_{K+1}\leqslant hk^{\frac{1}{2}\epsilon\tau}$ and, by (23),

$$\tfrac{1}{2}\epsilon\tau=\tfrac{1}{2}(n+1+1/\zeta)+\tfrac{1}{2}\epsilon-1<(1/\zeta)+\tfrac{1}{2}\epsilon-1, \tag{29}$$

we see that l satisfies the condition of lemma 3, and thus either (24) holds or, by (22),

$$|f(l)|>(\mathrm{e}^{2hk}\,c_6^{LR_{K+1}})^{-D}. \tag{30}$$

We show that the assumption that (30) is valid leads to a contradiction.

By (27), (28) and the inequalities

$$R_{K+1}\leqslant hk^{(1/\zeta)+\frac{1}{2}\epsilon-1}\leqslant H,\quad (n+1)\,\zeta<1,$$

we clearly have

$$|F(l)|<H^{R_K(S_{K+1}+1)}<\mathrm{e}^{\frac{1}{8}\delta H}$$

if H is sufficiently large. Further, since $LR_{K+1}\leqslant hk^{(1/\zeta)-\frac{1}{2}\epsilon}$, we deduce from (30) that

$$|f(l)|>2\,\mathrm{e}^{-\frac{1}{8}\delta H}.$$

Hence we obtain

$$|f(l)/F(l)|>2\,\mathrm{e}^{-\frac{1}{4}\delta H},$$

and, by virtue of the estimate for the double sum established above, it follows that the absolute value of the number on the right of (26) exceeds $\frac{1}{2}|f(l)/F(l)|$. Now let θ and Θ denote respectively the upper bound of $|f(z)|$ and the lower bound of $|F(z)|$ with z onΓ. Since $2|z-l|$, with z on Γ, exceeds the radius of Γ, we obtain from (26)

$$4\theta|F(l)|>\Theta|f(l)|. \tag{31}$$

It is clear that

$$\Theta\geqslant(\tfrac{1}{2}R_{K+1}\log k)^{R_K(S_{K+1}+1)},$$

and, by (21) of lemma 3, we have

$$\theta \leqslant \mathrm{e}^{2hk} c_5^{LR_{K+1}\log k}.$$

Thus from (28) we deduce that

$$\Theta\,|F(l)|^{-1} \geqslant (\tfrac{1}{2}\log k)^{R_K(S_{K+1}+1)},$$

and from (30) we obtain

$$\theta\,|f(l)|^{-1} \leqslant (\mathrm{e}^{2hk} c_5^{LR_{K+1}\log k})^{D+1}.$$

Then (31) gives

$$\log 4+(D+1)\{2hk+c_{17}LR_{K+1}\log k\} \geqslant R_K(S_{K+1}+1)\{\log\log k-\log 2\}. \tag{32}$$

But $LR_{K+1} \leqslant hk^{\frac{1}{2}\epsilon(K-1)+1}$ and so the number on the left of (32) is at most

$$c_{18}\,hk \quad \text{or} \quad c_{18}\,hk^{\frac{1}{2}\epsilon(K-1)+1}\log k$$

according as $K=0$ or $K>0$. On the other hand we have

$$R_K(S_{K+1}+1) \geqslant \tfrac{1}{2}hk^{\frac{1}{2}\epsilon K}(k/2^{K+1}),$$

and, since $K+1<\tau$, we see that the number on the right of (32) exceeds

$$c_{19}^{-1}\,hk^{\frac{1}{2}\epsilon K+1}\log\log k.$$

The two estimates are obviously inconsistent (both for $K=0$ and for $K>0$) if k is sufficiently large; the contradiction implies the validity of (24), and the lemma follows by induction.

Lemma 5. *For each integer j with $0 \leqslant j \leqslant k^n$ we have*

$$\log|\phi_j(0)| < -k^{\frac{1}{2}\epsilon(\tau-1)+1}/\log k, \tag{33}$$

where

$$\phi(z) = \Phi(z, \ldots, z).$$

Proof. We write, for brevity,

$$X = [k^{\frac{1}{2}\epsilon(\tau-1)}], \quad Y = [k/\log k].$$

Then clearly $[hk^{\frac{1}{2}\epsilon\sigma}] \geqslant X$ and $[k/2^\sigma] \geqslant Y$, where σ denotes the largest integer $<\tau$, and so, by lemma 4, we see that (15) holds for each integer l with $1 \leqslant l \leqslant X$ and each set of non-negative integers $m_1, \ldots, m_{n-1}$ satisfying $m_1+\ldots+m_{n-1} \leqslant Y$. Hence we have

$$|\phi_m(r)| \leqslant n^k\,\mathrm{e}^{-\frac{1}{2}\delta H} \tag{34}$$

for each integer r with $1 \leqslant r \leqslant X$ and each integer m satisfying $0 \leqslant m \leqslant Y$ (cf. the proof of lemma 4). Let Γ and Λ denote circles in the complex plane, described in the positive sense, with centres the origin and with radii $X\log k$ and $\frac{1}{4}$ respectively. Suppose further that w is any complex number on Λ. We proceed to calculate an upper bound for $|\phi(w)|$.

We write, for brevity,

$$E(z) = \{(z-1)\ldots(z-X)\}^{Y+1}.$$

By Cauchy's residue theorem we have (cf. lemma 4)

$$\frac{1}{2\pi\mathrm{i}}\int_\Gamma \frac{\phi(z)}{(z-w)\,E(z)}\,\mathrm{d}z = \frac{\phi(w)}{E(w)}+\frac{1}{2\pi\mathrm{i}}\sum_{r=1}^{X}\sum_{m=0}^{Y}\frac{\phi_m(r)}{m!}\int_{\Gamma_r}\frac{(z-r)^m\,\mathrm{d}z}{(z-w)\,E(z)}, \tag{35}$$

where, as before, Γ_r denotes the circle in the complex plane, described in the positive sense, with centre r and radius $\frac{1}{2}$. Since, for z on Γ_r,

$$|(z-r)^m/E(z)| < 8^{Y+1}$$

and since also, by (29),

$$X(Y+1) \leqslant k^{\frac{1}{2}\epsilon(\tau-1)+1} \leqslant k^{1/\zeta} \leqslant H,$$

it follows, on using (34), that the absolute value of the double sum on the right of (35) is at most

$$X(Y+1)\, 8^{Y+2} n^k \,\mathrm{e}^{-\frac{1}{2}\delta H} < (8n)^{k+2} H \,\mathrm{e}^{-\frac{1}{2}\delta H} < \mathrm{e}^{-\frac{1}{4}\delta H}.$$

Now let ξ and Ξ denote respectively the upper bound of $|\phi(z)|$ and the lower bound of $|E(z)|$ with z on Γ. From (35) we have

$$|\phi(w)| \leqslant \{2\xi\Xi^{-1} + \mathrm{e}^{-\frac{1}{4}\delta H}\}\, |E(w)|,$$

and it is clear that

$$|E(w)| \leqslant (X+1)^{X(Y+1)} \quad \text{and} \quad |\Xi| \geqslant (\tfrac{1}{2}X \log k)^{X(Y+1)}.$$

Since also, by (21) of lemma 3,

$$\xi \leqslant \mathrm{e}^{2hk}\, c_5^{LX \log k},$$

we obtain

$$|\phi(w)| \leqslant 2\,\mathrm{e}^{2hk}\, c_5^{LX \log k} (\tfrac{1}{4} \log k)^{-X(Y+1)} + (2X)^{2XY} \mathrm{e}^{-\frac{1}{4}\delta H},$$

and the second term on the right is at most $\mathrm{e}^{-\frac{1}{8}\delta H}$ since clearly, by (29),

$$2XY \log (2X) \leqslant 2\epsilon(\tau-1)\, k^{\frac{1}{2}\epsilon(\tau-1)+1} \leqslant \tfrac{1}{8}\delta H.$$

On noting that $L \leqslant k^{1-\epsilon}$ and

$$X(Y+1) \geqslant \tfrac{1}{2} k^{\frac{1}{2}\epsilon(\tau-1)+1}/\log k, \tag{36}$$

it follows easily that

$$|\phi(w)| \leqslant (\log k)^{-\frac{1}{2}X(Y+1)}.$$

Let now j be any integer satisfying $0 \leqslant j \leqslant k^n$. By Cauchy's residue theorem we have

$$\frac{j!}{2\pi \mathrm{i}} \int_\Lambda \frac{\phi(w)}{w^{j+1}}\, \mathrm{d}w = \phi_j(0).$$

Thus from the bound for $|\phi(w)|$ established above we obtain

$$|\phi_j(0)| < j!\, 4^j (\log k)^{-\frac{1}{2}X(Y+1)}.$$

Now clearly

$$j!\, 4^j \leqslant (4j)^j \leqslant k^{k^{n+1}},$$

and since, by (23),

$$\tfrac{1}{2}\epsilon(\tau-1)+1 = \tfrac{1}{2}(n+1+1/\zeta) > n+1, \tag{37}$$

we see that

$$\tfrac{1}{2}X(Y+1) \log\log k \geqslant \tfrac{1}{4} k^{\frac{1}{2}\epsilon(\tau-1)+1} \log\log k/\log k \geqslant 2k^{n+1} \log k.$$

Hence we have

$$|\phi_j(0)| \leqslant (\log k)^{-\frac{1}{4}X(Y+1)},$$

and this, together with (36), implies the validity of (33), as required.

Lemma 6. *Let $t_1, \ldots, t_n$ denote rational integers with absolute values at most T, and let*

$$W = t_1 \log \alpha_1 + \ldots + t_n \log \alpha_n.$$

Then either $W = 0$ or

$$|W| > c_{20}^{-T} A^{-2D|t_n|}.$$

Proof. Let $a_1, \dots, a_n$ be defined as in the proof of lemma 2 but with α_j^{-1} read for α_j if $t_j < 0$. Then

$$\omega = a_1^{|t_1|} \dots a_n^{|t_n|} (\alpha_1^{t_1} \dots \alpha_n^{t_n} - 1)$$

represents an algebraic integer with degree at most D. Further, by the first observation in § 3, we see that any of its conjugates, obtained by substituting arbitrary conjugates for $\alpha_1, \dots, \alpha_n$, has absolute value at most $c_{21}^T A^{2|t_n|}$. If $\omega = 0$ then W is a multiple of $2\pi\mathrm{i}$ and obviously the lemma is valid. Otherwise we have $|\mathrm{norm}\,\omega| \geqslant 1$ and hence

$$|\omega| \geqslant (c_{21}^T A^{2|t_n|})^{-D+1}.$$

But since

$$a_1^{|t_1|} \dots a_n^{|t_n|} \leqslant c_{22}^T A^{|t_n|},$$

we deduce from (12) that

$$|\omega| \leqslant |W|\, \mathrm{e}^{|W|} c_{22}^T A^{|t_n|},$$

and on assuming, as we may, that $|W| < 1$, the assertion follows.

5. Proof of theorem 4

We proceed to verify that the inequalities (33) cannot all be valid. This will suffice to establish theorem 4.

We write, for brevity,

$$R = (L_1+1) \dots (L_n+1) - 1.$$

Then any integer r with $0 \leqslant r \leqslant R$ can be expressed uniquely in the form

$$r = \lambda_1 + \lambda_2(L+1) + \dots + \lambda_n(L+1)^{n-1},$$

where $\lambda_1, \dots, \lambda_n$ denote integers satisfying $0 \leqslant \lambda_j \leqslant L_j$ $(1 \leqslant j \leqslant n)$. For each such r we define

$$p_r = p(\lambda_1, \dots, \lambda_n), \quad \psi_r = \lambda_1 \log \alpha_1 + \dots + \lambda_n \log \alpha_n,$$

and we write

$$\Psi_j = \sum_{r=0}^{R} p_r \psi_r^j \quad (0 \leqslant j \leqslant R).$$

Now it is easily seen that Ψ_j differs from $\phi_j(0)$ by at most an amount $\mathrm{e}^{-\frac{1}{2}\delta H}$. For clearly we have

$$\phi_j(0) = \sum_{\lambda_1=0}^{L_1} \dots \sum_{\lambda_n=0}^{L_n} p(\lambda_1, \dots, \lambda_n)\, (\gamma_1 \log \alpha_1 + \dots + \gamma_{n-1} \log \alpha_{n-1})^j,$$

and, since $|\psi_r| < c_{23} L$ and $j \leqslant R$, we obtain from (11)

$$|(\gamma_1 \log \alpha_1 + \dots + \gamma_{n-1} \log \alpha_{n-1})^j - \psi_r^j| < (c_{24} L)^R\, \mathrm{e}^{-\delta H};$$

this gives

$$|\phi_j(0) - \Psi_j| \leqslant (R+1)\, \mathrm{e}^{hk} (c_{24} L)^R\, \mathrm{e}^{-\delta H},$$

and the number on the right is at most $\mathrm{e}^{-\frac{1}{2}\delta H}$ since

$$L \leqslant k^{1-\epsilon}, \quad R \leqslant k^n \leqslant H^{\zeta n}, \quad \zeta < 1/(n+1).$$

It follows immediately from (33) (which is applicable since $j \leqslant k^n$), (29) and the inequality $k \leqslant H^{\zeta}$, that

$$\log |\Psi_j| < -\tfrac{1}{2} k^{\frac{1}{2}\epsilon(\tau-1)+1} / \log k. \tag{38}$$

We now observe that each integer p_r $(0 \leqslant r \leqslant R)$ satisfies an equation of the form

$$p_r \Delta_r(\psi_r) = \sum_{j=0}^{R} \sigma_j \Psi_j, \tag{39}$$

where $\Delta_r(x)$ and $\sigma_0, \ldots, \sigma_R$ are defined by the identities

$$\Delta_r(x) = \prod_{\substack{s=0\\ s\neq r}}^{R} (x-\psi_s) = \sigma_0 + \sigma_1 x + \ldots + \sigma_R x^R.$$

We shall compare estimates for the numbers on either side of (39). Clearly $\psi_r - \psi_s$ $(r \neq s)$ is a linear form in $\log \alpha_1, \ldots, \log \alpha_n$ with integer coefficients, each coefficient having absolute value at most $L \leqslant k^{1-\epsilon} \leqslant H$, and not all the coefficients vanishing. Thus by the hypothesis made at the outset, that the only integers $b'_1, \ldots, b'_n$ with absolute values at most H such that (9) holds are given by $b'_1 = \ldots = b'_n = 0$, we see that $\psi_r - \psi_s \neq 0$ and so, by lemma 6,

$$|\psi_r - \psi_s| > c_{20}^{-L} A^{-2D|L_n|}.$$

On noting that there are $R \leqslant k^n$ factors in the product defining Δ_r, and using also (18) and the inequality $L \leqslant k$, it follows easily that

$$\log |\Delta_r(\psi_r)| \geqslant -c_{25} k^{n+1}.$$

We proceed now to calculate an upper bound for $\log |\Delta_r(\psi_r)|$ for any r such that $p_r \neq 0$; there is certainly at least one r with this property. We have

$$|\sigma_j| \leqslant \prod_{\substack{s=0\\ s\neq r}}^{R} (1+|\psi_s|) \quad (0 \leqslant j \leqslant R),$$

and, by virtue of the trivial inequality $|\psi_j| \leqslant c_{26} k$, we see that the product on the right does not exceed $(c_{27} k)^R$. Hence from (38) and (39) we obtain

$$\log |p_r \Delta_r(\psi_r)| \leqslant \log (R+1) + R \log (c_{27} k) - \tfrac{1}{2} k^{\frac{1}{2}\epsilon(\tau-1)+1} / \log k.$$

By (37) and the inequality $R \leqslant k^n$, it follows that

$$\log |\Delta_r(\psi_r)| \leqslant -\tfrac{1}{4} k^{\frac{1}{2}\epsilon(\tau-1)+1} / \log k$$

if k is sufficiently large. But, by (37) again, we see that this is inconsistent with the lower bound for $\log |\Delta_r(\psi_r)|$ established above, and the contradiction proves the theorem.

6. Preliminaries to the proof of theorem 1

It remains now only to verify that theorem 3 implies the validity of theorem 1. The present section serves to supply the definitions and preliminary results which will be needed for the main argument given in §7.

We observe first that if α, β are algebraic numbers with degrees at most d and heights at most H then $\alpha+\beta$ and $\alpha\beta$ have degrees at most d^2 and heights at most H', where $\log H'/\log H$ is bounded above by a number depending only on d. For let a, b denote the leading coefficients in the minimal defining polynomials of α, β, and let $\alpha^{(i)}$, $\beta^{(j)}$ denote their respective conjugates. Then $\alpha+\beta$ and $\alpha\beta$ are zeros of the polynomials

$$(ab)^{d^2} \prod_{i,j} (x-(\alpha^{(i)}+\beta^{(j)})), \quad (ab)^{d^2} \prod_{i,j} (x-\alpha^{(i)}\beta^{(j)})$$

respectively, which clearly have integer coefficients and degrees at most d^2. The zeros of the minimal polynomials of $\alpha+\beta$ and $\alpha\beta$ are thus given by some subsets of the $\alpha^{(i)}+\beta^{(j)}$ and

$\alpha^{(i)}\beta^{(j)}$ respectively, and the leading coefficients divide $(ab)^{d^2}$. The assertion now follows from the fact that the $\alpha^{(i)}$, $\beta^{(j)}$ have absolute values at most dH. For later reference we observe that if α, α', β, β' $(\beta \neq \beta')$ are algebraic numbers with degrees at most d and heights at most H, then, by repeated application of the above results, it follows that $(\alpha-\alpha')/(\beta-\beta')$ has degree at most d^4 and height at most H', where, as before, $\log H'/\log H$ is bounded above by a number depending only on d; and, by the remark made at the beginning of § 3, the absolute value of $(\alpha-\alpha')/(\beta-\beta')$ is at most d^4H'.

Suppose now that $f(x,y)$, n, κ and m are defined as in § 1. We note immediately that, for the purpose of proving theorem 1, there is no loss of generality in assuming that the coefficient of x^n in $f(x,y)$ is ± 1. For let the coefficient be denoted by a. Then obviously $|a|^{n-1}f(x,y)$ can be expressed as a binary form $F(X,Y)$ with integer coefficients, where $X = ax$, $Y = y$; further, the coefficient of X^n in $F(X,Y)$ is ± 1. Now if $\kappa' = \frac{1}{2}(\kappa+n+1)$, so that $\kappa > \kappa' > n+1$, and if theorem 1 is valid with respect to F, then all solutions in integers X, Y of the equation $F(X,Y) = M$, where $M = m|a|^{n-1}$, satisfy

$$\max(|X|,|Y|) < C' \,\mathrm{e}^{(\log M)^{\kappa'}}$$

for some C' depending only on κ' and the coefficients of F; and the desired bound for $\max(|x|,|y|)$ follows easily.

The zeros of $f(x,1)$ will be denoted by $\alpha^{(1)},\ldots,\alpha^{(n)}$. It will be supposed that the arrangement is such that $\alpha^{(1)},\ldots,\alpha^{(s)}$ only are real, and that $\alpha^{(s+1)},\ldots,\alpha^{(s+t)}$ are the complex conjugates of $\alpha^{(s+t+1)},\ldots,\alpha^{(n)}$ respectively; thus it is implied that $n = s+2t$. The assumption made above, that the coefficient of x^n in $f(x,y)$ is ± 1, further implies that $\alpha^{(1)},\ldots,\alpha^{(n)}$ are algebraic integers. The algebraic number field generated by $\alpha = \alpha^{(1)}$ over the rationals will be denoted by K, and $\theta^{(1)},\ldots,\theta^{(n)}$ will represent the conjugates of any element θ of K corresponding to the conjugates $\alpha^{(1)},\ldots,\alpha^{(n)}$ of α. $C_1, C_2, \ldots$ will denote numbers, greater than 1, which can be specified explicitly in terms of n and the coefficients of f.

Finally, we shall denote by $\eta_1,\ldots,\eta_r$ a set of $r = s+t-1$ units in K such that the determinant Δ of order r with $\log|\eta_i^{(j)}|$ in the ith row and jth column does not vanish. That such a set of units exists is established, for example, by Landau (1918; p. 49, Satz 136; note here that $r > 0$ since $n \geqslant 3$) as an intermediate stage in the proof of a classical theorem of Dirichlet. It is not difficult to verify, by a study of Landau's exposition, for example, that the units $\eta_1,\ldots,\eta_r$ can further be chosen so that $|\Delta| > C_1^{-1}$ and

$$|\log|\eta_i^{(j)}|| < C_2 \quad (1 \leqslant i,j \leqslant r)$$

for some effectively computable numbers C_1, C_2 as above.

7. Proof of theorem 1

Suppose that x, y are rational integers satisfying (1), and put $\beta = x-\alpha y$. Clearly β is an algebraic integer in K, and we have

$$|\beta^{(1)}\ldots\beta^{(n)}| = m. \tag{40}$$

Further we assert that an associate γ of β can be determined such that

$$|\log(m^{-1/n}|\gamma^{(j)}|)| \leqslant C_3 \quad (1 \leqslant j \leqslant n) \tag{41}$$

for some number C_3 as defined in § 6. For, by the properties of $\eta_1,\ldots,\eta_r$ specified earlier,

every point P in r-dimensional Euclidean space occurs within a distance rC_2 of some point of the lattice with basis
$$(\log|\eta_i^{(1)}|, \ldots, \log|\eta_i^{(r)}|) \quad (1 \leqslant i \leqslant r).$$

On taking P as the point
$$(\log(m^{-1/n}|\beta^{(1)}|), \ldots, \log(m^{-1/n}|\beta^{(r)}|)),$$

it follows immediately that there exist rational integers $b_1, \ldots, b_r$ such that
$$|b_1 \log|\eta_1^{(j)}| + \ldots + b_r \log|\eta_r^{(j)}| + \log(m^{-1/n}|\beta^{(j)}|)| \leqslant rC_2 \quad (1 \leqslant j \leqslant r),$$

that is, such that
$$|\log(m^{-1/n}|\gamma^{(j)}|)| \leqslant rC_2 \quad (1 \leqslant j \leqslant r), \tag{42}$$

where
$$\gamma = \beta\eta_1^{b_1}\ldots\eta_r^{b_r}. \tag{43}$$

Since now $|\gamma^{(j+t)}| = |\gamma^{(j)}|$ $(s < j \leqslant s+t)$, we see that the inequalities (42) in fact hold for each j with $1 \leqslant j \leqslant n$, except possibly for $j = s+t$ and $j = s+2t$ (only one of which exists if $t = 0$). But by (40) and (43) we have
$$|\gamma^{(1)}\ldots\gamma^{(n)}| = m,$$

that is
$$\sum_{j=1}^{n} \log(m^{-1/n}|\gamma^{(j)}|) = 0,$$

and so the inequalities (41) must hold without exception for some C_3. Note that (41) implies in particular that the height of γ is at most $C_4 m$; for the leading coefficient in the minimal defining polynomial of γ is 1, all other coefficients can be expressed as elementary symmetric functions in $\gamma^{(1)}, \ldots, \gamma^{(n)}$, and, by (41), each of the latter has absolute value at most $e^{C_3} m^{1/n}$.

The equations
$$\log|\gamma^{(j)}/\beta^{(j)}| = b_1 \log|\eta_1^{(j)}| + \ldots + b_r \log|\eta_r^{(j)}| \quad (1 \leqslant j \leqslant r)$$

serve to express each number Δb_j as a linear combination of numbers $\log|\gamma^{(j)}/\beta^{(j)}|$ with coefficients given by certain cofactors of Δ. Thus, denoting by H the maximum of the absolute values of $b_1, \ldots, b_r$, and recalling the supposition that $|\Delta| > C_1^{-1}$, it follows easily that the maximum of the numbers
$$|\log|\gamma^{(j)}/\beta^{(j)}|| \quad (1 \leqslant j \leqslant r)$$

must exceed $C_5^{-1} H$ for some C_5 as above. Let this maximum be given by $j = J$. Then from (41) we have
$$|\log(m^{-1/n}|\beta^{(J)}|)| = |\log|\beta^{(J)}/\gamma^{(J)}| + \log(m^{-1/n}|\gamma^{(J)}|)| \geqslant C_5^{-1} H - C_3,$$

and since, by (40),
$$\sum_{j=1}^{n} \log(m^{-1/n}|\beta^{(j)}|) = 0,$$

it follows that
$$\log(m^{-1/n}|\beta^{(l)}|) \leqslant -(C_5^{-1} H - C_3)/(n-1) \tag{44}$$

for some superscript l. In particular we see that $|\beta^{(l)}| \leqslant C_6 m^{1/n}$ and so, again by (40), we have $|\beta^{(k)}| \geqslant C_6^{-1} m^{1/n}$ for some superscript $k \neq l$. Let j denote any superscript other than k or l (this exists since $n \geqslant 3$).

Now the identity
$$(\alpha^{(k)} - \alpha^{(l)})\beta^{(j)} - (\alpha^{(j)} - \alpha^{(l)})\beta^{(k)} = (\alpha^{(k)} - \alpha^{(j)})\beta^{(l)},$$

together with (43), gives

$$\alpha_1^{b_1}\ldots\alpha_r^{b_r}-\alpha_{r+1}=\omega, \tag{45}$$

where

$$\alpha_s=\eta_s^{(k)}/\eta_s^{(j)}\quad(1\leqslant s\leqslant r),$$

$$\alpha_{r+1}=\frac{(\alpha^{(j)}-\alpha^{(l)})\,\gamma^{(k)}}{(\alpha^{(k)}-\alpha^{(l)})\,\gamma^{(j)}}\quad\text{and}\quad\omega=\frac{(\alpha^{(k)}-\alpha^{(j)})\,\beta^{(l)}\gamma^{(k)}}{(\alpha^{(k)}-\alpha^{(l)})\,\beta^{(k)}\gamma^{(j)}}.$$

By (44) and the choice of k we see that

$$|\beta^{(l)}/\beta^{(k)}|\leqslant C_7\,\mathrm{e}^{-H/C_8}. \tag{46}$$

Further, by the properties of $\eta_1,\ldots,\eta_r$ specified in § 6 and by the preliminary observations made in that section, $\alpha_1,\ldots,\alpha_r$ represent algebraic numbers with degrees at most n^2 and with heights not exceeding some number C_9 as above. Furthermore, again by the results of § 6, we deduce easily that α_{r+1} is an algebraic number with degree at most n^6 and with height at most $(C_{10}m)^{C_{11}}$ (where C_{11} in fact depends only on n). Also, by virtue of (41), we have

$$\max\,(|\alpha_{r+1}|,|\alpha_{r+1}|^{-1})\leqslant C_{12},$$

and we obtain similarly

$$0<|\omega|<C_{13}\,|\beta^{(l)}/\beta^{(k)}|. \tag{47}$$

It is clear that (45), (46) and (47) imply the validity of (4) with n replaced by $r+1$ and with $\delta=\frac{1}{2}C_8^{-1}$, provided that $H>2C_7\,C_8\,C_{13}$.

We apply theorem 3 with α_s, n and δ defined as above, with $d=n^6$ and with κ given by $\kappa'=\frac{1}{2}(\kappa+n+1)$; this choice for κ is in accordance with the hypotheses since clearly $\kappa'>n+1$, $r\leqslant n-1$ and $r<n-1$ if $\alpha_1,\ldots,\alpha_r$ are not all real. We conclude that

$$H<\max\,(C',\{C_{11}\log\,(C_{10}m)\}^{\kappa'}) \tag{48}$$

for some effectively computable number C' depending only on† n, κ and the coefficients of f. But now the identities

$$x=(\alpha^{(2)}\beta^{(1)}-\alpha^{(1)}\beta^{(2)})/(\alpha^{(2)}-\alpha^{(1)}),\quad y=(\beta^{(1)}-\beta^{(2)})/(\alpha^{(2)}-\alpha^{(1)})$$

imply that

$$\max\,(|x|,|y|)<C_{14}\max\,(|\beta^{(1)}|,|\beta^{(2)}|),$$

and by (42) and (43) we obtain

$$|\beta^{(j)}|=|\gamma^{(j)}\eta_1^{(j)-b_1}\ldots\eta_r^{(j)-b_r}|\leqslant m^{1/n}C_{15}^H\quad(1\leqslant j\leqslant n).$$

Hence (48) gives

$$\max\,(|x|,|y|)<m^{1/n}\max\,(C'',C_{16}^{\{\log(C_{10}m)\}^{\kappa'}})$$

for some $C''=C''(n,\kappa,f)$. Since $\kappa>\kappa'>n+1$, the number on the right is certainly less than $C\,\mathrm{e}^{(\log m)^\kappa}$ for a suitable C, and this completes the proof of the theorem.

References

Baker, A. 1964*a* Rational approximations to certain algebraic numbers. *Proc. Lond. Math. Soc.* **41**, 385–398.

Baker, A. 1964*b* Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers. *Quart. J. Math. Oxford* (2) **15**, 375–383.

Baker, A. 1966 Linear forms in the logarithms of algebraic numbers. *Mathematika* **13**, 204–216.

† Note here that it is not necessary to make explicit reference to r, since it follows as a consequence of the wording of theorem 3 that, in the expression for C, one can substitute for n any integer n' satisfying $n'\geqslant n$ and $\kappa>n'+1$ or $\kappa>n'+2$ according as $\alpha_1,\ldots,\alpha_n$ are or are not all real.

Baker, A. 1967 Linear forms in the logarithms of algebraic numbers (II). *Mathematika* **14**, 102–107.

Borevich, Z. I. & Shafarevich, I. R. 1966 *Number theory*. New York and London: Academic Press.

Dyson, F. J. 1947 The approximation to algebraic numbers by rationals. *Acta Math.* **79**, 225–240.

Gelfond, A. O. 1952 *Transcendental and algebraic numbers*. Moscow (in Russian); English translation 1960. New York: Dover Publications.

Hilbert, D. 1901 Mathematische Probleme. *Arch. Math. Phys.* **1**, 44–63, 213–237; = *Ges. Abhandlungen* III, 290–329.

Landau, E. 1918 *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale*. Leipzig and Berlin: Teubner.

Liouville, J. 1844 Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques. *Comptes rendus* **18**, 883–885, 910–911; *J. Math. pures appl.* **16** (1851), 133–142.

Mordell, L. J. 1913 The Diophantine equation $y^2-k = x^3$. *Proc. Lond. Math. Soc.* (2) **13**, 60–80.

Mordell, L. J. 1947 *A chapter in the theory of numbers*. Cambridge University Press.

Mordell, L. J. 1963 The Diophantine equation $y^2 = ax^3+bx^2+cx+d$, or fifty years after. *J. Lond. Math. Soc.* **38**, 454–458.

Roth, K. F. 1955 Rational approximations to algebraic numbers. *Mathematika* **2**, 1–20.

Schneider, Th. 1936 Über die Approximation algebraischer Zahlen. *J. Reine Angew. Math.* **175**, 182–192.

Schneider, Th. 1957 *Einführung in die transzendenten Zahlen*. Berlin, Göttingen, Heidelberg: Springer.

Siegel, C. L. 1921 Approximation algebraischer Zahlen. *Math. Z.* **10**, 173–213; = *Ges. Abhandlungen* I, 6–46.

Siegel, C. L. 1929 Über einige Anwendungen diophantischer Approximationen. *Abh. Preuss. Akad. Wiss.* No. 1; = *Ges. Abhandlungen* I, 209–266.

Skolem, Th. 1935 Einige Sätze über $\mathfrak{p}$-adische Potenzreihen mit Anwendung auf gewisse exponentielle Gleichungen. *Math. Ann.* **111**, 399–424.

Skolem, Th. 1938 *Diophantische Gleichungen. Ergebnisse Math.* 5 (4). Springer: Berlin. Reprinted 1950, Chelsea: New York.

Thue, A. 1909 Über Annäherungswerte algebraischer Zahlen. *J. Reine Angew. Math.* **135**, 284–305.