(62) in which \( h = [h_1, \ldots, h_l] \). Therefore, by Lemma 12 there exists a vector \( h^0 = [h^0_1, \ldots, h^0_l] \) such that the system of equations

\[
M_{h^0}(a_1, a_1, \ldots, a_k) = 0 \quad (i \in I)
\]

is solvable in integers. Denoting a solution by \( [x^0_1, x^0_2, \ldots, x^0_k] \) we get from (60) and (61) for all \( i \leq l \)

\[
\begin{align*}
\sum_{j=1}^{k} a^0_{ij} x^0_j + b^0_{ij} &= 0, \\
\sum_{j=1}^{k} a^0_{ij} x^0_j - b^0_{ij} &= 0 \quad (1 \leq s \leq r)
\end{align*}
\]

hence by (59)

\[
\prod_{i=1}^{l} \left( \prod_{j=1}^{k} a^0_{ij} - b_{ij} \right) = 0 \quad (1 \leq i \leq l).
\]

References


Correction to [6]

p. 401; insert after formula (8):

provided \( p_i > 2, \ a \equiv 1 \mod p_i \) or \( p_i = 2, \ a \equiv 1 \mod 4 \).

Received on 16. 10. 1975

Diophantine approximation in power series fields*

by

WOLFGANG M. SCHMIDT (Boulder, Colo.)

Dedicated to Professor Theodor Schneider on his 65th birthday

1. Introduction

1.1. The setting. Let \( K \) be the field of formal series

\[
a = a_k x^k + a_{k-1} x^{k-1} + \ldots
\]

with an arbitrary integer \( k \) and with coefficients in a given field \( F \) of characteristic zero. The rational functions in \( x \) with coefficients in \( F \) form a subfield \( K_0 = F(x) \) of \( K \), and the polynomials form a subring \( S = F[x] \). In \( K \) we have the non-archimedean valuation with

\[
|a| = 2^k
\]

if the leading coefficient in (1) is \( a_k \neq 0 \). If \( f \) is a polynomial, then

\[
|f| = 2^{|\deg f|}.
\]

Many results on "ordinary" diophantine approximation, i.e. approximation of reals by rationals, carry over to approximation of elements of \( K \) by rational functions, i.e. by elements of \( K_0 \).

For example, Dirichlet's Theorem holds: If \( ax \) does not lie in \( K_0 \), then there are infinitely many rational functions \( f/g = f(x)/g(x) \) in \( K_0 \)

\[
|a - (f/g)| \leq |g|^{-2}.
\]

Also Liouville's Theorem holds: If \( ax \) is algebraic over \( K_0 \) of degree \( s > 1 \), then for every rational function \( f/g \), we have

\[
|a - (f/g)| > c_s |g|^{-s},
\]

with a constant \( c_s(a) > 0 \). Now just as in ordinary diophantine approxi-
mation, where Liouville's Theorem was eventually strengthened to Roth's Theorem [5], so in the power series case, Liouville's Theorem was strengthened by Uehiya [7] to the following result, which we shall call the Roth–Uehiya Theorem: If \( a \) is algebraic of degree \( \alpha > 1 \), then for \( \varepsilon > 0 \) and for every rational function \( f/g \), we have

\[
|a - (f/g)| > c_{\varepsilon}(a, \varepsilon)|g|^{-\alpha - \varepsilon},
\]

with a constant \( c_{\varepsilon}(a, \varepsilon) > 0 \).

As was first observed by Kolchin [1], new questions arise in the power series case in connection with algebraic differential equations. These questions have no analog in ordinary diophantine approximation. Denote the formal derivatives of a series \( a \) by \( a^{(0)}, a^{(1)}, \ldots \). Suppose \( a \in K_\alpha \) satisfies an algebraic differential equation, i.e., an equation

\[
A(a, a^{(0)}, \ldots, a^{(\alpha)}) = 0,
\]

where \( A(Y, Y_1, \ldots, Y_l) \) is a non-zero polynomial in variables \( Y, Y_1, \ldots, Y_l \) with coefficients in \( K_\alpha \). Then Kolchin proved that for rational functions \( f/g \),

\[
|a - (f/g)| > c_{\varepsilon}(a)|g|^{-d},
\]

where \( c_{\varepsilon}(a) > 0 \) and where \( d \) is the \( \alpha \)-denomination of \( A \), i.e., the maximum of

\[
a_0 + 2a_1 + \ldots + (l+1)a_l
\]

over all monomials \( Y_{a_0}Y_{a_1}^2 \cdots Y_{a_l}^l \) occurring in \( A \) with non-zero coefficients. When \( l = 0 \), i.e., when \( a \) is algebraic, then (4) reduces to Liouville's estimate (2). It is not known whether the exponent \( -d \) in (4) is best possible if \( l > 0 \); perhaps the exponent should always be \( -(2 + \varepsilon) \) for \( \varepsilon > 0 \), or even \(-2\). Contributions to this question were made by Osgood [3], and special equations were discussed by Schmidt [6]. Osgood [2] used Kolchin's Theorem to prove a result on algebraic functions which is stronger than Liouville's estimate (2) and weaker* than the Roth–Uehiya estimate (3), but which in contrast to the latter is effective in some sense.

Now in ordinary diophantine approximation one deals not only with approximation by rationals, but also with approximation by algebraic numbers. Similarly, in the power series case, one deals with approximation by algebraic functions. In the present paper we intend to discuss approximation by solutions of algebraic differential equations. In particular, we shall prove generalizations of Liouville's Theorem, dealing with approximation to a solution of an algebraic differential equation by solutions of other differential equations.

* Except that for functions of degree 3 it is stronger.

Of course, the "expansions about infinity" (1) could be replaced by other expansions, say by expansions \( a_0 Y + a_{k+1} Y^{k+1} + \ldots \) about zero. The expansions (1) have the advantage that now the polynomials play the role which is played by the integers in ordinary diophantine approximation.

1.2. Heights. Let \( Y \) be a "differential variable", i.e., a variable with which we associate further variables \( Y_1, Y_2, \ldots \), representing "derivatives" of \( Y \). A differential monomial will be an expression

\[
P(Y) = P(Y, Y_1, \ldots) = Y^{a_0}Y_1^{a_1} \cdots Y_l^{a_l}
\]

with nonnegative integers \( a_0, \ldots, a_l \). We write \( K[Y] \) (or \( S[Y] \)) for the ring of differential polynomials

\[
A = \sum_{n} \pi_n P_n,
\]

where the sum is finite, where the \( P_n \) are differential monomials and where the coefficients \( \pi_n \) lie in \( K \) (or in \( S \)). Given a differential polynomial

\[
f = \sum_{n} f_n P_n
\]

with \( f_n \in S = F[X] \) and with distinct monomials \( P_n \), define its height \( \mathcal{H}(A) \) by

\[
\mathcal{H}(A) = \max_n \mathcal{H}(f)
\]

Let \( A(a) \) be obtained by substituting \( a, a^{(0)}, a^{(2)}, \ldots \) for \( Y, Y_1, Y_2, \ldots \) into \( A(Y) \).

Let \( \Omega(m) \) be the set of \( \beta \in K \) which satisfy a linear differential equation of order \( \leq m \):

\[
B(\beta) = 0,
\]

where

\[
B = g_m Y_m + \cdots + g_1 Y_1 + g_0 Y + f
\]

is a non-zero linear differential polynomial of order \( \leq m \) with coefficients \( g_m, \ldots, f \) in \( S = F[X] \). Here \( \mathcal{H}(B) = \max \{|g_m|, \ldots, |g_0|, |f| \} \). Given \( \beta \in \Omega(m) \), the height \( \mathcal{H}(\beta) \) is given by

\[
\mathcal{H}(\beta) = \min_{B \in \Omega(m)} \mathcal{H}(B),
\]

with the minimum to be taken over all non-zero linear differential polynomials \( B \in S[Y] \) of order \( \leq m \) with \( B(\beta) = 0 \).

If \( P \) is again given by (5), define the order \( l(P) \) by

\[
l(P) = \begin{cases} l & \text{if } a_l > 0, \\ -1 & \text{if } a_0 = \ldots = a_l = 0. \end{cases}
\]
Also put
\[ s(P) = a_0 + \cdots + a_l, \]
\[ r(P) = a_0 + 3a_1 + \cdots +(2l-1)a_l. \]
If \( A \) is a differential polynomial, write \( l(A) \), \( s(A) \), \( r(A) \), respectively, for the maximum of \( l(P) \), \( s(P) \), \( r(P) \) over the monomials \( P \) occurring in \( A \) with non-zero coefficients. Then \( l(A) \) is the order of \( A \).

Let \( \Omega(n, s) \) be the set of \( \beta \in \mathbb{R} \) which satisfy a non-trivial differential equation \( B(\beta) = 0 \) with
\[ \exists \in \mathbb{R} \]
\[ s(B) \leq m, \quad s(B) \leq s. \]
Given \( \beta \in \Omega(n, s) \), the height \( S_{\mathbb{R}}(\beta) \) is given by
\[ S_{\mathbb{R}}(\beta) = \min_B S(B), \]
with the minimum to be taken over all non-zero \( B \) with \( (9) \) and with \( B(\beta) = 0 \).

1.3. The results. First about solutions of linear differential equations:

**Theorem 1.** Suppose \( \alpha \in \Omega(l) \), but \( \not\in \Omega(l-1) \). Then there is \( a_0(\alpha) > 0 \) such that
\[ |\alpha - \beta| > a_0(\alpha)S(\beta)^{-1}, \]
for every \( \beta \in \Omega(l) \).

In general we have

**Theorem 2.** Suppose \( m \geq 0, \ s \geq 1 \) are given. Suppose \( \alpha \) satisfies a non-trivial differential equation \( A(\alpha) = 0 \), but satisfies no non-trivial differential equation of order \( < m \). Then for \( \beta \neq \alpha \), \( \beta \in \Omega(m, s) \), we have
\[ |\alpha - \beta| > c_0S(\beta)^{-1}, \]
where \( c_0 = a_0(m, s, \epsilon) > 0 \) and where
\[ c_0 = 3m^{m+1}r(lA)^m(m+1). \]

No special importance attaches to our value of the constant \( c_0 \).

The restriction that a satisfies no differential equation of order \( < m \) could be removed at the cost of further complication and a possible change of our value for \( c_0 \). Since \( \Omega(0, 1) \) consists of rational functions, the case \( m = 0, \ s = 1 \) is Kolchin’s Theorem, except for the just mentioned restriction except for the value of \( c_0 \).

Both Theorems 1 and 2 are Liouville type results. We have as yet no Roth type results.

The reader will not be surprised to hear that resultants play an essential role in the proof. More surprising is the fact that existence theorems on power series solutions of differential equations are needed.

2. Differential equations (1)

2.1. Notation. Suppose \( A = A(Y) = A(Y, Y_1, \ldots, Y_s) \) is a differential polynomial. The derivative \( A(1) \) of \( A \) is defined by
\[ A(1) = \frac{\partial A}{\partial Y} + \frac{\partial A}{\partial Y_1}Y_1 + \frac{\partial A}{\partial Y_2}Y_2 + \cdots + \frac{\partial A}{\partial Y_s}Y_s. \]

All the usual rules on differentiation of sums and products hold. The higher derivatives \( A(2), A(3), \ldots \) are defined by induction. It is easily seen that for \( \alpha \in \mathbb{C} \) we have
\[ A(1)(\alpha) = (A(\alpha))^{(1)}. \]

Given two differential polynomials \( A, B \), put
\[ A \circ B = A(B, B^{(1)}, \ldots). \]

Then \( (A \circ B) \circ C = A \circ (B \circ C) \), and in particular \( (A \circ B)(\alpha) = A(B(\alpha)) \).

In the introduction we used the multiplicative valuation \( |a| \), in order to stress the analogy with ordinary diophantine approximation. In what follows, it will be more convenient to use the additive function \( \psi \) with \( \psi(0) = -\infty \) and
\[ \psi(a) = k \]
if \( a \) is given by \( (1) \) with \( \epsilon_0 \neq 0 \). Clearly \( |a| = 2^{\psi(a)} \). We have \( \psi(a + \beta) \leq \max\{\psi(a), \psi(\beta)\} \), and equality holds here if \( \psi(a) = \psi(\beta) \). Not \( \psi \) itself, but \( -\psi \) is an ‘additive valuation’.

2.2. Linear polynomials. Given a homogeneous linear differential polynomial
\[ L = \lambda_1 Y + \lambda_2 Y_1 + \cdots + \lambda_s Y_s, \]
put \( w(L) = -\infty \) if \( L = 0 \) and
\[ w(L) = \max_{0 < i < 2} \psi(\lambda_i) - i \]
otherwise. Then certainly \( w(L_1 + L_2) \leq \max\{w(L_1), w(L_2)\} \), with equality if \( w(L_1) = w(L_2) \). Given a non-zero homogeneous linear differential polynomial \( L \) with \( w(L) = w_0 \), we may write
\[ L = L_{w_0} + L_{w_0 - 1} + \cdots, \]
where
\[ L_j = a_{j_0} X^{j_0} Y + a_{j_1} X^{j_1+1} Y_1 + \cdots + a_{j_s} X^{j_s+s} Y_s \]
with constant coefficients, and where \( L_{w_0} \neq 0 \).

---

(1) The lemmas of this section are not new, but are collected here for convenience. See e.g. [4].
Substituting $a = X^l$, we obtain

$$L_j(X^l) = (a_0 + a_1 t + a_2 t(t-1) + \cdots + a_i t(t-1)^i \cdots (t-i+1)) X^{l+i}$$

say, with certain polynomials $p_j$ of degree $\leq l$. The polynomial $p_j = p_{L_t}$ is not zero and is the indicial polynomial of $L$. The indicial polynomial of $L = 0$ is identically zero. It follows from (12), (14) that

$$w(L) = \max_i \{v(L(X^i)) - v(X^{l+i})\},$$

hence that

$$w(L) = \max_a \{v(L(a)) - v(a)\},$$

and that in fact

$$w(L) = v(L(a)) - v(a)$$

if $p_L(a) \neq 0$.

**Lemma 1.** Let $L$, $M$ be homogeneous linear differential polynomials. Then $M \circ L$ is a homogeneous linear differential polynomial with

$$l(M \circ L) = l(M) + l(L),$$

$$w(M \circ L) = w(M) + w(L),$$

and

$$p_{M \circ L}(t) = p_M(t) p_L(t + w(L)).$$

**Proof.** We may suppose that $L$, $M$ are non-zero. Only (19), (20) require a proof. By (12), (14),

$$L(X^i) = p_L(t) X^{i+u(L)} + p_{L \circ L - 1}(t) X^{i+u(L)-1} + \cdots$$

and

$$M(X^i) = p_M(t) X^{i+u(M)} + p_{M \circ M - 1}(t) X^{i+u(M)-1} + \cdots$$

Thus

$$(M \circ L)(X^i) = p_M(t) p_L(t + w(L)) X^{i+u(L)+u(M)} + \cdots,$$

and (19), (20) become obvious.

2.3. Solutions of differential equations. Every differential polynomial $A$ may uniquely be written as

$$A = \lambda + L + \overline{A},$$

where $\lambda \in K$, where $L$ is homogeneous and linear, and $\overline{A}$ is a sum

$$\overline{A} = \sum a_u P_u,$$

with differential monomials $P_u$ having $s(P_u) \geq 2$. Given

$$a_u P_u = a_u X^n X^n t \cdots Y^n t,$$

write $w(a_u P_u) = v(a_u) - a_1 - 3a_2 - \cdots - la_l$, and set $w(\overline{A})$ for the maximum of $w(a_u P_u)$ over all summations in (22), with the understanding that $w(\overline{A}) = - \infty$ if $\overline{A} = 0$. Now (recall the definition of $s(P)$ in §1.2),

$$v(a_u P_u(X^i)) \leq s(P_u) + w(a_u P_u),$$

and if $t \leq 0$, $s(P_u) \geq 2$, then

$$v(\pi u P_u(X^i)) \leq 2t + w(a_u P_u) = 2v(X^i) + w(a_u P_u).$$

Thus if $t \leq 0$, then

$$v(A(X^i)) \leq 2v(X^i) + w(A),$$

and, more generally, if $v(a) \leq 0$, then

$$v(A(a)) \leq 2v(a) + w(A).$$

**Lemma 2.** Suppose $A$ is of the type (21) with $\lambda = 0$, $L \neq 0$. Suppose $\eta \in K$, $\eta \neq 0$, is a solution of the differential equation

$$A(\eta) = 0.$$

Suppose that either $A = 0$, or that $A \neq 0$ and

$$v(\eta) < \min(0, w(L) - w(A)).$$

Then $p_L(v(\eta)) = 0$.

**Proof.** Suppose we had $p_L(v(\eta)) \neq 0$. If $A = 0$, then by (17),

$$-\infty = v(0) = v(L(\eta)) = v(\eta) + w(L) \neq -\infty$$

which is impossible. If $A \neq 0$ and (24) holds, then $L(\eta) + A(\eta) = 0$, whence by (17), (23),

$$v(\eta) + w(L) = v(L(\eta)) = v(A(\eta)) \leq 2v(\eta) + w(A) < v(\eta) + w(L),$$

which again is impossible.

2.4. Existence of solutions.

**Lemma 3.** Suppose $L$ is a non-zero homogeneous linear differential polynomial with $w(L) = w$. For a fixed $t$, consider series

$$\eta = a_t X^t + a_{t-1} X^{t-1} + \cdots$$

with undetermined coefficients. Then

$$L(\eta) = b_t X^t + b_{t-1} X^{t-1} + \cdots$$

with $b_{t+u-j} = p_L(t-j) a_t + b_{t+j-u}$, where $b_{t+u-j}$ is a linear form in $a_t, a_{t-1}, \ldots, a_{t+j}$. (In particular, $b_{t+0} = 0$.)
proof. It suffices to observe that
(a) every coefficient in the series \( L(a_t x^t + \cdots + a_{t+j} x^{t+j}) \) is a linear form in \( a_t, \ldots, a_{t+j} \).
(b) \( L(a_t x^t + a_{t+1} x^{t+1} + \cdots) = P_L(t-j) a_{t+j} x^{t+w-j} + \cdots \).

**Lemma 4.** Suppose \( \bar{A} \) is a differential polynomial with summands \( \pi_i P_i \) having \( \sigma(P_i) \geq 2 \). Let \( t, w \) be given with
\[
t < \min \left( 0, w - w(\bar{A}) \right).
\]
Consider series (23) with undetermined coefficients. Then
\[
\bar{A}(\eta) = a_{t+w} x^{t+w} + a_{t+w+1} x^{t+w+1} + \cdots,
\]
where \( a_{t+w} \) is a polynomial in \( a_t, \ldots, a_{t+j} \) with constant term zero. (In particular, \( a_{t+w} = 0 \).

**Proof.** We may suppose that \( \bar{A} \) is of the special type
\[
\bar{A} = x^s Y_1 \cdots Y_s
\]
with \( w(\bar{A}) = -1 - 1 - \cdots - 1 \). Then the coefficients on the right hand side of (27) are forms in \( a_t, a_{t+1}, \ldots \) of degree \( s \). A typical summand of \( \bar{A}(\eta) \) will be some constant times
\[
a_{t+j_1} \cdots a_{t+j_m} x^{w(\eta) - j_1 - \cdots - j_m}.
\]
The exponent here will equal \( t+w-j \) if
\[
t + w - j = st - j_1 - \cdots - j_m - w(\bar{A}).
\]
Then by (26),
\[
j_k < j_1 + \cdots + j_m = (s-1)t + w(\bar{A}) - w = t + w(\bar{A}) - w + j < j.
\]
So \( a_{t+w} \) is a polynomial in \( a_t, \ldots, a_{t+j} \).

**Lemma 5.** Suppose \( \bar{A} \) is given by (21), with \( L, \lambda \not= 0 \). Suppose that \( \pi_L(t) \) has no integer root \( 0 \leq \nu(\lambda) - w(L) \). Suppose that either \( \bar{A} = 0 \) or that \( \bar{A} \not= 0 \) and
\[
\nu(\lambda) < \min \{ w(L), 2w(L) - w(\bar{A}) \}.
\]
Then there is an \( \eta \in \mathbb{K} \) with \( \nu(\eta) = \nu(\lambda) - w(L) \) having
\[
\bar{A}(\eta) = 0.
\]
**Proof.** Put \( w = w(L) \), \( \nu = \nu(\lambda) \), \( t = \nu - w \). Write \( \eta \) in the form (25).

Then
\[
\nu(\bar{A}) = \bar{A}(\eta) = d_0 x^\nu + \cdots + d_{\nu-1} x^{\nu-1} + \cdots,
\]
with coefficients
\[
d_{\nu-j} = \pi_L(t-j) a_{t+j} + \cdots,
\]
where \( d_{\nu-j} \) is a polynomial in \( a_t, \ldots, a_{t+j} \). This follows from Lemma 3 if \( \bar{A} = 0 \), and it follows from Lemmas 3, 4 if \( \bar{A} \not= 0 \). In the latter case observe that the condition (28) is satisfied in view of (28). By our hypothesis on \( \pi_L \), the coefficient \( p_L(t-j) a_{t+j} \) of \( a_{t+j} \) is non-zero for \( j = 0, 1, \ldots \) We thus can successively choose \( a_t, a_{t+1}, \ldots \) such that \( L(\eta) + \bar{A}(\eta) = 0 \), i.e. that \( A(\eta) = 0 \). Since \( v(\lambda) = \nu \), we have \( d_0 = p_L(\nu) a_{\nu-j} \), whence \( a_t \not= 0 \), and \( v(\eta) = t = \nu - w \).

3. Approximation by solutions of linear differential equations

3.1. Linear differential ideals. The linear differential polynomials
\[
A = \lambda + L = \lambda + \lambda_1 Y_1 + \ldots + \lambda_l Y_l
\]
form a vector space \( V \) over \( \mathbb{K} \). Given \( l \geq -1 \), the linear differential polynomials of order \( \leq l \) form a subspace \( V_l \) of \( V \) of dimension \( l+2 \) with basis \( 1, Y, \ldots, Y_l \). A linear differential ideal is defined as a subspace of \( V \) which is closed under taking derivatives and which is equal to \( V \) if it contains a non-zero element of order \( -1 \). The principal ideal \( (A) \) generated by \( A \not= 0 \) is the intersection of the linear differential ideals containing \( A \). If \( I(\lambda) \geq 0 \), then \( (A) \) is the subspace of \( V \) spanned by \( A, A^2, \ldots \). Hence \( (A) \) consists of the polynomials \( L \circ A \), where \( L \) is a homogeneous linear differential polynomial. If \( I(\lambda) = -1 \), then \( (A) = V \).

**Lemma 6.** Every non-zero linear differential ideal is a principal ideal.

**Proof.** Given a non-zero ideal \( I \), let \( A \) be a non-zero element of \( I \) of least order; say \( I(A) = l \). If \( l = -1 \) then \( \mathbb{I} = V \) and thus \( I = (A) \). So suppose that \( l \geq 0 \). Let \( B \) be an arbitrary non-zero element of \( I \) with \( I(B) = m \). The polynomials
\[
B, A, A^2, \ldots, A^{m-1}
\]
span a subspace \( S \) of \( V \). Since \( A \) with \( I(A) = l \) is a non-zero element of \( I \) with least possible order, the intersection \( \cap \{ V_{-l} \} = 0 \). So \( \dim S + \dim V_{l-1} = \dim S + V_{l-1} = \dim V_{l-1} = \dim V_{m-l} - 1 = m - l + 1 \). We conclude that the vectors \( (29) \) are linearly dependent. Since \( A, A^2, \ldots \) are linearly independent, \( B \) must be a linear combination of \( A, A^2, \ldots, A^{m-1} \). Therefore \( B \in (A) \) and \( I = (A) \).

If \( (A) = (B) \), then clearly \( I(A) = I(B) \). Thus the order \( I(\lambda) \) of a linear differential ideal \( I \) may be defined as the order of any of its generators.

Write \( (A, B) \) for the ideal generated by \( A, B \).

3.2. Differential resultants. Suppose \( A \not= 0 \), \( B \not= 0 \) generate a differential ideal \( I \). Put \( I = I(A) = I(B) \), \( r = I(\lambda) \), and suppose that
\[
r < l, \quad r < m.
\]
If
\[ A = \lambda + \lambda_1 Y + \ldots + \lambda_t Y_t, \]
then
\[ A^{(i)} = \lambda_1^{(i)} + \lambda_2^{(i)} Y + \ldots + \lambda_{t+1}^{(i)} Y_{t+1} + \lambda_{t+1} Y_{t+1} \quad (i = 1, 2, \ldots), \]
where the \( \lambda_1^{(i)} \) and the \( \lambda_i^{(i)} \) are certain linear combinations of \( \lambda \) and the \( \lambda_i \)'s and their derivatives. Similarly, if
\[ B = \mu + \mu_1 Y + \ldots + \mu_m Y_m, \]
then
\[ B^{(i)} = \mu_1^{(i)} + \mu_2^{(i)} Y + \ldots + \mu_{m+1}^{(i)} Y_{m+1} + \mu_{m+1} Y_{m+1} \quad (i = 1, 2, \ldots). \]
The determinant \( R(Y) \) with \( l + m - 2t \) rows and columns, given by
\[
R(Y) = \begin{vmatrix}
A(Y) & \lambda_2 & \ldots & \lambda_t \\
A^{(1)}(Y) & \lambda_1^{(1)} & \ldots & \lambda_1^{(t)} \\
\vdots & \vdots & \ddots & \vdots \\
A^{(m-t)}(Y) & \lambda_{m-t+1}^{(m-t-1)} & \ldots & \lambda_t^{(m-t-1)} \\
B(Y) & \mu_2 & \ldots & \mu_t \\
B^{(1)}(Y) & \mu_1^{(1)} & \ldots & \mu_1^{(t)} \\
\vdots & \vdots & \ddots & \vdots \\
B^{(m-t)}(Y) & \mu_{m-t+1}^{(m-t-1)} & \ldots & \mu_t^{(m-t-1)}
\end{vmatrix}
\]
is a linear differential polynomial called the resultant of \( A \) and \( B \).

Lemma 7. The resultant \( R(Y) \) is a non-zero linear differential polynomial of order \( r \). The ideal \( \mathcal{I} = (A, B) \) is generated by \( R \): \( \mathcal{I} = (R) \).

Proof. The resultant is a linear combination of
\[
A, A^{(1)}, \ldots, A^{(m-t-1)}, B, B^{(1)}, \ldots, B^{(m-t-1)},
\]
hence belongs to \( \mathcal{I} \). Being a linear combination of the polynomials (31), it is a linear combination of \( 1, Y, \ldots, Y_{t+m-1} \). But from our definition of \( R(Y) \) as a determinant and from elementary linear algebra it follows that the coefficients of \( Y_{t+1}, \ldots, Y_{t+m-1} \) are in fact zero, so that \( \mathcal{I}(R) \triangleq \mathcal{R} \) is of order \( r \). It will thus suffice to show that \( \mathcal{R} \neq 0 \); since \( \mathcal{I} \subset \mathcal{I}(R) \), this will guarantee that \( \mathcal{I}(R) = r \) and that in fact \( \mathcal{I} = (R) \).

Now for \( s \geq l, m \), the \( 2s-l-m+2 \) polynomials
\[
A, A^{(1)}, \ldots, A^{(s-1)}, B, B^{(1)}, \ldots, B^{(s-m)}
\]
lie in the vector space \( V_s \) of dimension \( s+2 \). So for large \( s \), the vectors (32) will be linearly dependent; let \( s \) be the smallest integer with this property. We may then suppose that \( A^{(s-1)} \) is a linear combination of
\[
A, A^{(1)}, \ldots, A^{(s-1)}, B, B^{(1)}, \ldots, B^{(s-m)}, \]
are linearly independent and span \( \mathcal{I} = (A, B) \). In fact, if \( t \) is large, then \( \mathcal{I} \cap V_t \) is spanned by
\[
A, A^{(1)}, \ldots, A^{(s-1)}, B, B^{(1)}, \ldots, B^{(s-m)},
\]
so that \( \dim (\mathcal{I} \cap V_t) = 2s-l-m+1 \). On the other hand \( \mathcal{I} = (C) \) for some \( C \) with \( l(C) = r \), so that \( \mathcal{I} \cap V_t \) is spanned by \( C, C^{(1)}, \ldots, C^{(r-1)} \) if \( r \geq 0 \), and by \( 1, Y, \ldots, Y_t \) if \( r = -1 \), and \( \dim (\mathcal{I} \cap V_t) = t+1-r \). Comparing our formulae we find that
\[
s = l+m-r.
\]

The polynomials
\[
A, A^{(1)}, \ldots, A^{(s-1)}, B, B^{(1)}, \ldots, B^{(s-m)}
\]
are linearly independent by our minimal choice of \( s \). They span a subspace of \( V_{s-1} \) of dimension \( 2s-l-m = s-r \), while \( V_r \) is of dimension \( r+2 \). Now \( (s-r)+(r+2) = s+2 > \dim V_{s-1} \). So the subspace spanned by (34) has a non-zero intersection with \( V_r \); there exists a non-zero linear combination \( D \) of (34) which lies in \( V_r \). Since \( D \), of course, is a vector subspace of \( \mathcal{I} \), it is unique except for a factor \( \gamma \neq 0 \). Since the polynomials (34) are linearly independent, there are coefficients \( a, a_1, \ldots, a_{s-1} \) and \( \beta, \beta_1, \ldots, \beta_{s-1} \), not all zero, and unique except for a common factor \( \gamma \neq 0 \), such that
\[
aA + a_1 A^{(1)} + \ldots + a_{s-1} A^{(s-1)} + \beta B + \beta_1 B^{(1)} + \ldots + \beta_{s-1} B^{(s-m-1)}
\]
lies in \( V_r \). Now since \( s-l-m-1 = r-1 \) and \( s-m-l = r-1 \), this uniqueness means precisely that the submatrix obtained from the determinant for \( R(Y) \) by crossing out the first column, is of rank \( l+m-2s-1 \). So in \( E \), which is a linear combination of (31), some coefficient of this linear combination is non-zero. Moreover, the polynomials (31) are the same as (34); hence they are linearly independent. So \( R \neq 0 \).

3.3. Proof of Theorem 1. Introduce the additive height
\[
h_A = \max \{ v(a) \}
\]
of a differential polynomial \( A \) given by (7). Then \( S(A) = 2^m A \). Define \( h_m(\beta) \) and \( h_{n+1}(\beta) \) in the obvious way. Theorem 1 now says that if \( a \) lies in \( \mathcal{O}(l) \) but not in \( \mathcal{O}(l-1) \), then
\[
v(a-\beta) > -(l+1) h_l(\beta) - c_l(a)
\]
for every \( \beta \neq a \) in \( \mathcal{O}(l) \).
Since $a$ lies in $\mathcal{O}(l)$, it satisfies an equation $A(a) = 0$ with a non-zero linear differential polynomial $A \in \mathcal{S}(Y)$ having $l(A) \leq l$. Since $a$ does not lie in $\mathcal{O}(l-1)$, we have in fact $l(A) = l$, and $A$ is unique except for a non-zero factor in $X_0$. Say $A = \lambda + L$ with $L \neq 0$.

Suppose at first that $A(\beta) = 0$. Then $\eta = a - \beta$ satisfies $L(\eta) = 0$, and $p_L(\eta(a)) = 0$ by Lemma 2. So this case is impossible if $p_L$ has no integer root, and $v(a-\beta) = v(\eta) \geq t_0$ where $t_0$ is the least integer root of $p_L$ if there are such integer roots.

By Lemma 2, we may therefore suppose that $\lambda \neq 0$. The element $\beta \in \mathcal{O}(1)$ satisfies a non-trivial linear differential equation $B(\beta) = 0$ with $B \in \mathcal{S}(Y)$, with $m = l(B) \leq l$ and with $h(B) = h_1(\beta)$. Since $A(\beta) \neq 0$, the polynomials $A, B$ are not proportional, and $B(a) \neq 0$. Now if, say, $B = \mu + M$, then $B(a) = B(a) - B(\beta) = M(a-\beta)$ and
\[
v(B[a]) \leq v(a-\beta) + h(M) \leq v(a-\beta) + h(B).
\]
So (35) will follow if we can prove that
\[
v(B(a)) > -il(B) - c'(a).
\]
Hence it remains to prove
\[v(a-\beta) + h(B) + (m-r)h(A).
\]
This follows from $A(a) = 0$ and the determinant formula for $R$ if $r < m$, and is trivial if $r = m$. Now since $r \geq -1$ and since $(m-r)h(A) \leq c'(a)$, Theorem 1 will follow from
\[v(B(a)) > \begin{cases} 0 & \text{if } p_L \text{ has no integer roots,} \\ \min(0, t_0 - l) & \text{if } t_0 \text{ is the least integer root of } p_L. 
\end{cases}
\]
To prove Lemma 8, we observe that $R$ is a "constant", i.e. a polynomial in $\mathcal{S}(X)$, if $r = -1$. Then $v(R(a)) = v(R) \geq 0$. We may thus suppose that $r \geq 0$. Write
\[R = r_0 + \cdots + r_n Y + r_{n+1} Y_1 + \cdots + r_{n+j} Y_j,
\]
with $N \neq 0$. Since $A \in \mathcal{S}(R)$, there is a homogeneous linear differential $Q \in X[Y]$ with
\[A = Q \circ R, \quad \text{whence with } \quad L = Q \circ N.
\]
Consider the following linear differential equation in $\eta$:
\[
\begin{align*}
N(\eta) + B(a) &= 0, \\
N(\eta) + B(a) &= (Q \circ N(\eta) + (Q \circ B)(a) = 0.
\end{align*}
\]
By Lemma 2, such an $\eta$ has $p_L(\eta(a)) = 0$.

Observe that by Lemma 1, the indicial polynomial $p_L$ of $N$ is a divisor of the indicial polynomial $p_L$ of $L$. If $p_L$ has no integer roots, then neither does $p_L$. Then (36) has a solution $\eta$ by (37).

Theorem 1 has $p_L(\eta(a)) = 0$, which is impossible. So the case $r \geq 0$ is impossible if $p_L$ has no integer roots.

If $p_L$ has integer roots and if $v(R(a)) < t_0 - l$, then $p_L$ and hence $p_L$ has no integer root $\leq v(R(a)) + l$. Since $N \in \mathcal{S}(Y)$, we have $w(N) \geq -l(N) = -r > -l$, and $p_L$ and $p_L$ have no integer roots $\leq v(R(a)) - w(N)$. Again by Lemma 5, the equation (38) has a solution $\eta$ by (37). Since again $p_L(\eta(a)) = 0$ by (39), we have again a contradiction.

4. Approximation by solutions of general differential equations

4.1. A reduction. Let $m, s, a$ be as in Theorem 2. So $a$ satisfies a non-trivial equation $A(a) = 0$ with $A \in \mathcal{S}(Y)$ and with $l(A) \geq m$, but it satisfies no such equation of an order less than $m$. We may suppose that $A$ is the minimal possible value of $v(A)$ (the functional $r$ is defined in §1.2), and therefore $(\partial A/\partial Y_i)(a) \neq 0$, since $r(\partial A/\partial Y_i) < r(A)$.

The inequality (1) of Theorem 2 which is to be proved says that
\[
v(a-\beta) > -c \min(0, -r_0 - m) - c
\]
for $\beta \in \mathcal{O}(m, s)$ distinct from $a$. (The constants $c$ here and in the sequel are not necessarily equal, and they will depend only on $a, m, s$.)

Let us first suppose that $A(\beta) = 0$. Then $\eta = \beta - a$ has $A(\eta + \beta - a) = 0$, or $C(\eta) = 0$, where $C \in X[Y]$ is given by
\[C(\eta) = A(a + \beta).
\]
Writing $C = \lambda + L + \odot$ as in (31), we have $\lambda = A(a) = 0$ and
\[
L = \left(\frac{\partial A}{\partial Y_1}(a) Y_1 + \cdots + \frac{\partial A}{\partial Y_n}(a) Y_n + \right) + \left(\frac{\partial A}{\partial Y_1}(\eta) Y_1 + \cdots + \frac{\partial A}{\partial Y_n}(\eta) Y_n\right).
\]
Observe that $L$ depends only on $a$ and that the coefficient of $Y_1$ is non-zero. By Lemma 2, we have either $p_L(\eta(a)) = 0$ or $C \neq 0$ and $v(\eta) \geq \min(0, w(L) - w(\odot))$. So certainly $v(\eta) \geq -c$, or $v(a-\beta) \geq -c$. 

---

6 — Acta Arithmetica XXXII.3
Now a may or it may not satisfy a differential equation $A_n(a) = 0$ of order $m$. If so, we may suppose $A_n$ to be an irreducible polynomial in $F[X, Y, \ldots, Y_m]$. Since certainly $a_1, \ldots, a_m$ are algebraically independent over $F(X) = K_0$, every polynomial $\hat{A}_n \in F[X, Y, \ldots, Y_m]$ with $A_n(a) = 0$ would have to be a multiple of $A_n$. One sees as above that $v(a - \beta) \geq -c$ if $A_n(\beta) = 0$ and $\beta \neq a$.

$v(\beta) = v(a)$ satisfies a non-trivial equation $B(\beta) = 0$ with $B \in S[Y]$, $l(B) \leq m$, $s(B) \leq s$, and $h(B) = h_{max}(\beta)$. Since $s(B_2X_1B_2) = s(B_2) + s(B_2)$ and $h(B_2X_1B_2) = h(B_2) + h(B_2)$, we may suppose $B$ to be an irreducible polynomial in $F[X, Y, \ldots, Y_m]$. Now $B(a) = 0$ is only possible if the polynomial $A_n$ above really exists and $B, A_n$ are proportional. Then $A_n(\beta) = 0$, and $v(a - \beta) \geq -c$ by what we said above. We may thus assume that $B(a) \neq 0$.

In order to prove (40), we may clearly suppose that $v(a - \beta) < v(a)$, whence that

$$v(\beta) = v(a).$$

If a monomial $P$ is given by (5), then $P(a) - P(\beta)$ is a sum of terms

$$(\beta^{l_0})^a \cdots (\beta^{l_k})^{a_k} \cdots (\beta^{l_{k-1}})^{a_{k-1}} + \cdots + (\beta^{l_0})^{a_k} \cdots (\beta^{l_k})^{a_k} \cdots (\beta^{l_{k-1}})^{a_{k-1}} \cdots (\beta^{l_0})^{a_{k-1}}$$

so that

$$v(P(a) - P(\beta)) \leq v(a - \beta) + v(a)(a_k + \cdots + a_{k-1}) - 1.$$ 

Hence

$$v(B(a)) = v(B(a) - B(\beta)) \leq v(a - \beta) + h(B) + c.$$ 

Clearly (40) would follow from

$$v(B(a)) \geq -(c_k - 1)h(B) - c.$$ 

Hence Theorem 2 is reduced to

**Theorem 2'.** Suppose $m \geq 0$, $s \geq 1$ are given. Suppose $a$ satisfies a non-trivial equation $A_n(a) = 0$ with $l = l(A) \geq m$, but satisfies no equation of order $< m$. Then every $B \in S[Y]$ with $l(B) \leq m, s(B) \leq s$, and with $B(a) \neq 0$ has

$$v(B(a)) \geq -(c_k - 1)h(B) - c$$

with

$$c_k = (9^{m+1} - 1)[r(A)s]^m - 1 \leq c_k - 1.$$ 

In fact, Theorem 2 for a particular value of $m$ follows from Theorem 2' with that particular value of $m$.

Theorem 2' will be proved by induction on $m$. Now if $B \in S[Y]$ has $l(B) = -1$ and $B(a) \neq 0$, then it is a non-zero constant, i.e., a non-zero polynomial in $X$, hence has $v(B(a)) = v(B) \geq 0$. Hence Theorem 2' is true for $m = -1$, with $c_k$ replaced by 0. Thus our inductive argument is off the ground, and we may suppose the theorem to be true for $m - 1$.

**4.2. An application of resultants.** Let $B = B(Y, \ldots, Y_m)$ and $E = E(Y, \ldots, Y_m)$ be differential polynomials in $S[Y]$ of respective degrees $b > 0$ and $e > 0$. Clearly $b \leq s(B)$ and $e \leq s(E)$. Suppose that $B$ and $E$ have no common factor of positive degree in $Y, \ldots, Y_m$. Then if we interpret them as polynomials in $Y, \ldots, Y_m$ with coefficients in the field $F(X, Y, \ldots, Y_m)$, they have no common factor of positive degree in $Y, \ldots, Y_m$. Hence their resultant $R = R(Y, \ldots, Y_m)$ will not be 0. (In contrast to § 3, we are now dealing with an ordinary resultant, not a differential resultant.) The following facts follow from the theory of resultants:

(45) $s(E) \leq s(B) + s(E) - eb \leq s(B)e.$

(Namely, each of the $(b + e)!$ summands in the determinant of order $b + e$ for $R$ can be estimated in this way.)

(46) $h(E) \leq h(B)s(E) + h(E)b \leq h(B)s(E) + h(B)e.$

There exist polynomials $U(Y, \ldots, Y_m)$ and $V(Y, \ldots, Y_m)$ in $S[Y]$ with

$$R = UB + VE,$$

having

(47) $s(U) \leq s(B)s(E) - s(B)$, $s(V) \leq s(E) - s(B)$,

(48) $h(U) \leq h(B)s(E) + h(E) - h(B)$,

(49) $h(V) \leq h(E)s(B) + h(EB) - h(B)$.

**Lemma 9.** Let $B \in S[Y]$ be given with $l(B) = m, s(B) \leq s$, and irreducible as a polynomial in $F(X, Y, \ldots, Y_m)$. Put

$$B_n = \partial B \partial X_m.$$ 

Then if $a$ is as in Theorem 2', we have

(50) $v(B_n(a)) > -c_k h(B) - c$ with

$$c_k = 2 \cdot 9^{m+1} s[r(A)s]^{m-1} = 2sc_k[r(A), m - 1, e^2].$$

Proof. Let $E$ be the resultant of $B$ and $E = B_n$. Applying the estimates on the resultant just given, we obtain

(51) $s(E) \leq s^2$, $h(E) \leq 2sh(B)$, and $R = UB + VB_n$ with

$$s(U), s(V) \leq c,$$

$$h(U), h(V) \leq (2s - 1)h(B) < 2sh(B).$$
It follows that $v(B(a))$ and $v(B(a))$ are $< 2s(B) + c$, whence that
\begin{equation}
\label{54}
v(B(a)) \leq \max(v(B(a)), v(B_{n}(a))) + 2s(B) + c.
\end{equation}
Since $l(B) < m$, our inductive hypothesis yields
\begin{align*}
v(B(a)) &> -c_{1}[r(A), m-1, s(B)]h(B) - c \\
&> -c_{1}[r(A), m-1, s^{2}]h(B) - c,
\end{align*}
whence by (53), (54),
\begin{align*}
\max[v(B(a)), v(B_{n}(a))] &\geq -\left(1 + c_{2}[r(A), m-1, s^{2}]\right) 2s(B) h(B) - c \\
&\geq -c_{2}[r(A), m-1, s^{2}] 2s(B) h(B) - c \\
&\geq -c_{2}h(B) - c.
\end{align*}

In our proof of Theorem 2' we may always suppose that $l(B) = m$ and that $B$ is irreducible. For both $h(B)$ and $v(B(a))$ are additive functions of $B$, i.e., $h(B_{n}(a)) = h(B_{n}) + h(B_{n})$ and $v(B_{n}(a)) = v(B(a)) + v(B_{n}(a))$. Since Theorem 2' is certainly true if $v(B(a)) > -c_{2}h(B) - c$, we may suppose that
\begin{equation}
\label{55}
v(B_{n}(a)) > -c_{2}h(B) - c.
\end{equation}

4.3. Manipulations with polynomials. Let $m \geq 0$ be fixed. If $P$ is a differential monomial given by (5), put $r_{m}(P) = 0$ if $l \leq m$, and $r_{m}(P) = \alpha_{m+1} + \alpha_{m+2} + \ldots + (2l-2m-1)\alpha_{l}$ otherwise. If $A$ is a differential polynomial given by (8) with non-zero coefficients $A_{n}$, write $r_{m}(A)$ for the maximum of $r_{m}(P_{n})$. Write $r_{m}(A) = 0$ if $A = 0$. Observe that $r_{m}(A) \leq r(A)$ and that
\begin{equation}
\label{56}
r_{m}(A^{(0)}) \leq r_{m}(A) + 2.
\end{equation}

If $B \in S[Y]$ with $l(B) = m$, then as in § 4.2 put $B_{n} = \partial B/\partial Y_{n}$. Then $h(B_{n}) \leq h(B)$ and $s(B_{n}) \leq s(B) - 1$.

**Lemma 10.** Given $B \in S[Y]$ with $l(B) = m$, we have
\begin{equation}
\label{57}
B^{(0)} = H_{1}(Y_{1}, \ldots, Y_{m+1}) + B_{n} Y_{m+1} \quad (j = 1, 2, \ldots),
\end{equation}
where $H_{j} \in S[Y]$ with
\begin{align}
\label{58}
s(H_{j}) &\leq s(B), \quad h(H_{j}) \leq h(B), \\
r_{m}(H_{j}) &\leq 2 - 2j.
\end{align}

**Proof.** Observe that $B^{(0)} = H_{1}(Y_{1}, \ldots, Y_{m}) + B_{n} Y_{m+1}$ with
\begin{equation}
\label{59}
H_{1} = \frac{\partial B}{\partial X} Y_{1} + \frac{\partial B}{\partial Y} Y_{1} + \ldots + \frac{\partial B}{\partial Y_{m-1}} Y_{m}.
\end{equation}
Clearly $s(H_{1}) \leq s(B)$ and $h(H_{1}) \leq h(B)$. Since $Y_{m+1}, \ldots$ do not occur in $H_{1}$, we have $r_{m}(H_{j}) = 0$. Finally, since $H_{1} \in S[Y]$, the lemma is true for $j = 1$. Assuming its truth for $j$, we have
\begin{equation}
\label{60}
B^{(j+1)} = H_{j+1}(Y_{1}, \ldots, Y_{m+1}) + B_{n} Y_{m+1} + 2 - 2j.
\end{equation}

\begin{equation}
\label{61}
H_{j+1} = H^{(j+1)} + B_{n} Y_{m+1},
\end{equation}

and
\begin{equation}
\label{62}
r_{m}(H_{j+1}) \leq \max(r_{m}(H_{j}) + 2, r_{m}(B_{n}) + 2j) \leq 2j.
\end{equation}

Since $H_{j+1} \in S[Y]$, the lemma is established for $j+1$.

**Lemma 11.** Let $m, n, l$ be integers with $m \geq 0$, $n \geq 1$ and $m + n = l$. Let $B$ with $l(B) = m$ and $A = A(Y_{1}, \ldots, Y_{l})$ be differential polynomials in $S[Y]$, and let $r \geq r_{m}(A)$. Then
\begin{equation}
\label{63}
B^{(r)} = D(Y_{1}, \ldots, Y_{m}) + C_{1}(Y_{1}, \ldots, Y_{m+1}) + \ldots + C_{n}(Y_{1}, \ldots, Y_{l}) B^{(n)},
\end{equation}
with polynomials $D, C_{1}, \ldots, C_{n}$ in $S[Y]$ having

\begin{equation}
\label{64}
s(D) \leq r(s(B)) + s(A) \quad (= s_{0}, \text{e.g.}),
\end{equation}

\begin{equation}
\label{65}
h(D) \leq r(h(B)) + h(A) \quad (= h_{0}, \text{e.g.}).
\end{equation}

**Proof.** We proceed by induction on $n$, and note that a degenerate version of the lemma is true for $n = 0$. Let $\mathcal{W}(s, h)$ be the set of expressions of the type of the right hand side of (58), with $D, C_{1}, \ldots, C_{n}$ in $S[Y]$ having
\begin{align*}
\max(s(D), s(C_{1} B^{(0)}), \ldots, s(C_{n} B^{(n)})) &\leq s, \\
\max(h(D), h(C_{1} B^{(0)}), \ldots, h(C_{n} B^{(n)})) &\leq h.
\end{align*}

Then we have to show that $B^{(r)} A$ lies in $\mathcal{W}(s_{0}, h_{0})$. Since $\mathcal{W}(s_{0}, h_{0})$ is closed under addition, it will suffice to show that for every "monomial"
\begin{equation}
\label{66}
M = C(Y_{1}, \ldots, Y_{m}) Y_{m+1}^{m+1} \ldots Y_{l}^{q},
\end{equation}
with $C \in S[Y]$ having $h(C) \leq h(A)$ and $s(C) \leq s(A) - \alpha_{m+1} - \ldots - \alpha_{l}$, and with $r_{m}(M) = \alpha_{m+1} + \ldots + (2l-2m-1)\alpha_{l} \leq r$, the product $B^{(r)} M$ lies in $\mathcal{W}(s_{0}, h_{0})$. Observe that
\begin{equation}
\label{67}
B^{(r)} M = B^{(r)} - r_{m}(M) C \sum_{j=1}^{n} (B_{a} Y_{m+1}^{m+1})^{\alpha_{m+1}}.
\end{equation}
Clearly

\[ B^*_{m+1} \equiv \mathbb{H}(s(B) - 1) + h(B) \]

We shall show that for \( i = 1, \ldots, n \),

\[ B^*_{m+1} Y_{m+i} \equiv \mathbb{H}(2i-1)(s(B) - 1) + (2i-1)h(B) \]

Now since

\[ \mathbb{H}(s_1, h_1) \mathbb{H}(s_2, h_2) \equiv \mathbb{H}(s_1 + s_2, h_1 + h_2) \]

i.e. the product of elements in \( \mathbb{H}(s_1, h_1) \) and in \( \mathbb{H}(s_2, h_2) \) lies in \( \mathbb{H}(s_1 + s_2, h_1 + h_2) \), it will follow from (61), (62) that \( B^*_{m+1} \equiv \mathbb{H}(s(B), h(B)) \). If \( n > 1 \), then (62) for \( i = 1, \ldots, n - 1 \) is true by induction. It will therefore suffice to prove (62) for \( i = n \). By Lemma 10,

\[ B^*_{m+1} Y_{m+n} = B^*_{m+1} B^{n+1} - B^*_{m+1} H_3(Y, \ldots, Y_{m+n-1}) \]

The first summand on the right lies in

\[ \mathbb{H}(2n-1, s(B) - 1) + (2n-1)h(B) \]

But so does the second, by virtue of \( r_m(H_3) \leq 2n-2 \) and our induction on \( n \). The lemma is proved.

4.4. Two cases. As in \S 4.1, let \( \alpha \) satisfy \( A(\alpha) = 0 \) with \( A \neq 0 \), \( A \neq \mathbb{S}[Y] \), and let it satisfy no equation of order less than \( m \). Since \( s(B_2) = s(B_1) + s(B_3) \) and \( h(B_2) = h(B_1) + h(B_3) \), it will suffice to prove Theorem 4 for irreducible polynomials \( B \) with \( l(B) = m \). In what follows, \( B \) will be such a polynomial. As usual, put \( l = l(A) \).

Now either

(i) \( m < l \), so that \( l = m+n \) with \( n \geq 1 \). Then set \( r = r(A) \) and construct \( D(Y) = D(Y, \ldots, Y_m) \) as in Lemma 11. Or

(ii) \( m = l \). Then put \( D = A \).

In case (i), the polynomials \( C_1 \) of Lemma 11 have \( s(C_1) \leq s \) and \( h(C_1) \leq r(A)h(B) + h(A) \leq r(A)h(B) + c \), so that \( v(C_1) \leq r(A)h(B) + c \). Now since \( A(\alpha) = 0 \) and \( A(\beta) = 0 \) yields

\[ v(D(\alpha)) \leq v(B(\alpha)) + r(A)h(B) + c \]

We also note that by (69), (60),

\[ s(D) \leq r(A)s(B) \]

Both (63) and (64) are trivially true in case (ii).

Again we distinguish two cases. Either \( B \nmid D \), i.e. \( B = B(X, Y, \ldots, Y_m) \) does not divide \( D = D(X, Y, \ldots, Y_m) \). Or \( B \mid D \).

4.5. The case \( B \nmid D \). Since \( B \) was irreducible, the polynomials \( B, D \) have no common factor. We form the resultant of \( B \) and \( E = D \), as explained in \S 4.2. The equation (47) becomes

\[ R = UB + VD \]

The polynomials \( U, V \) have

\[ s(U) \leq c, \quad s(V) \leq c \]

by (48), (64), and

\[ h(U) \leq h(B)s(D) + h(D)s(B) \leq 2r(A)s(B)h(B) + c \]

by (49), (64). We have \( v(U(a)) \leq h(U) + c \) and \( v(V(a)) \leq h(V) + c \)

by (66). So by (63), (65), (67), (68), we obtain

\[ v(B(a)) \leq \max\{v(U(a)) + v(B(a)), v(V(a)) + v(D(a))\} \]

\[ \leq v(B(a)) + 3r(A)s(B)h(B) + c \]

Also note that

\[ s(R) \leq s(B)s(D) \leq r(A)s^2(B) \]

by (45), (64), and that

\[ h(R) \leq 2r(A)s(B)h(B) + c \]

by (46), (64).

Now \( I(R) \leq m. \) We may therefore apply our induction hypothesis to \( R \) and we obtain

\[ v(B(a)) \geq -c_2h(B) - c_3 \]

with \( c_2 = c_2(r(A), m, -1, v(B)) \). In conjunction with (69), (71), we obtain

\[ v(B(a)) \geq -c_2h(B) - c_3 \]

with

\[ c_2 = 2r(A)s(B)c_2 + 3r(A)s(B) \leq c_2(r(A), m, s) \]

4.6. The case \( B \mid D \). We now have \( D = C_0(Y, \ldots, Y_mB) \), so that (58) becomes

\[ B^*A = C_0(Y, \ldots, Y_mB + C_0(Y, \ldots, Y_mB)B^{n+1} + \ldots + C_0(Y, \ldots, Y_mB)^{n+1} \]

This is certainly true in case (i), i.e. when \( m < l \). In case (ii) we have \( D_A \), so now \( B \) divides \( A \), and (72) reduces to \( B^*A = C_0B \). In the polynomial identity (72) we substitute \( a + Y \) for \( Y \), i.e. \( a + Y, \alpha^{(0)} + Y, \ldots \) for \( Y, Y_1, \ldots \). We obtain

\[ B^*(a+Y)A(a+Y) \]

\[ = C_0(a+Y)B(a+Y) + \ldots + C_0(a+Y)B^{(n)}(a+Y) \]

The differential polynomials on both sides of this equation are in \( K[Y] \) but not necessarily in \( S[Y] \).
Write the polynomial on either side of (73) as
\[ v + N + \overline{E}, \]
with the same conventions as in (21). Since \( A(a) = 0 \), we have \( v = 0 \).
Moreover, \( N \) is \( B'_*(a) \) times the linear part of \( A(a+y) \). If we write, as in (41), (42),
\[ A(a+y) = L + \overline{c} = \left[ \frac{\partial A}{\partial y}(a) \right] Y + \ldots + \left[ \frac{\partial A}{\partial y}(a) \right] Y_1 + \overline{c}, \]
then
\[ N = B'_*(a)L. \]

Thus \( N \neq 0 \), since \( L \neq 0 \) and since \( B_*(a) \neq 0 \) by (55).

In the notation (21), write
\[ C_i(a+y) = C_i(a) + L_i + \overline{c}_i, \]
\[ B(a+y) = B(a) + L_B + \overline{B}, \]
where of course \( C_i(a), B(a) \) are constants (i.e. elements of \( K \)), where \( L_i, L_B \) are linear, and \( C_i, B \) contain the non-linear terms. Now \( B^0(a+y) = (B(a+y))^0 \), so that
\[ B^0(a+y) = B^0(a) + L_B^0 + \overline{B^0}. \]

Further \( N \), being the linear term of (73), may be written as
\[ N = N_1 + N_2, \]
with
\[ N_1 = C_0(a)L_0 + C_1(a)L_1^0 + \ldots + C_n(a)L_n^0, \]
\[ N_2 = B(a)L_0 + B^0(a)L_1 + \ldots + B^{(n)}(a)L_n. \]

In view of (74), \( w(N) = w(L) + w(B) \). (The functional \( w \) was defined in § 2.2.) Here \( L \) and therefore \( w(L) \) depends only on \( a \), and \( v(B(a)) \) may be estimated by (55). So
\[ w(N) = -r(A)c_0 \overline{h}(B) - c. \]

On the other hand,
\[ \overline{L}_i = \left[ \frac{\partial C_i}{\partial y}(a) \right] Y + \ldots + \left[ \frac{\partial C_i}{\partial y}(a) \right] Y_1. \]

By (50), (60), we have \( v(C_i) \leq c \) and \( h(C_i) \leq r(A)h(B) + c \), whence
\[ v(\partial C_i \partial Y_i(a)) \leq r(A)h(B) + c, \]
and
\[ w(L_i) \leq r(A)h(B) + c. \]

So by (77),
\[ w(N_2) \leq v(B(a)) + r(A)h(B) + c. \]

Now either \( w(N_2) \geq w(N) \). Then by (78),
\[ v(B(a)) > -r(A)(c_0 + 1)h(B) - c \geq -c_0 \overline{h}(B) - c, \]
and (43) is true.

Or \( w(N_2) < w(N) \). Then by (75),
\[ w(N_1) = w(N), \]
and \( N, N_1 \) have the same indicial polynomial \( p_N \). Looking at (76) we see that
\[ N_1 = Q \circ L_B. \]

with \( Q = C_0(a)Y + \ldots + C_n(a)Y_n \). So by Lemma 1, the indicial polynomial \( p_{L_B} \) is a divisor of the indicial polynomial \( p_N \) of \( N_1 \). Moreover, \( v(Q) \leq r(A)h(B) + c \), whence \( w(Q) \geq r(A)h(B) + c \), and
\[ w(L_B) = w(N_1) - w(Q) \geq w(N) - r(A)h(B) - c. \]

4.7. Solving a differential equation. Consider the following differential equation for \( \eta \):
\[ B(a + \eta) = 0, \]
or
\[ B(a) + L_B(\eta) + B(\eta) = 0. \]

Now \( w(B) \leq h(B) + c \), so that
\[ \min = \min(w(L_B), 2w(L_B) - w(B)) \]
has
\[ \min > -(3c_0 + 3)r(A)h(B) - c \]
by (78), (79). Now if \( \min \leq v(B(a)) \), then (43) of Theorem 2 holds. If not, then condition (28) of Lemma 5 holds. So there is a solution \( \eta \) of (80) with \( v(\eta) = v(B(a)) - w(L_B) \). Putting \( \beta = a + \eta \) we have
\[ B(\beta) = 0 \]
and, by (78), (79),
\[ v(\beta - \beta) = v(\eta) \leq v(B(a)) + (c_0 + 1)r(A)h(B) + c. \]

Now obviously the right hand side of (73) vanishes if we substitute \( \beta \) and hence either \( A(\beta) = 0 \) or \( B_*(\beta) = 0 \). If \( A(\beta) = 0 \), then \( v(\beta - \beta) \geq -c \) by what we said in § 4.1, and (81) implies the desired (43) of Theorem 2.

There remains the case \( B_*(\beta) = 0 \). Then \( E(\beta) = 0 \), where \( E \) is the resultant of \( B, B_\ast \) constructed in the proof of Lemma 9. Since \( l(R) < m_0 \), it follows from the case \( m-1 \) of Theorem 2 (which follows from the case...
Large values of Dirichlet polynomials, IV

by

M. N. HUXLEY (Cardiff) and M. JUTILIA (Turku)

1. Introduction. This paper continues \[4\textendash }6\], [8], [9]. Our object is to estimate the size of a set of pairs \((s, \chi)\) at which a Dirichlet polynomial \(F(s, \chi)\) can be large. A precise statement is given in the next section where the notation is introduced. Our main tool is the reflection argument of \[4\], which we use in a simplified form due to Jutila [8], [9] as Lemma 8 below. It relates Dirichlet polynomials of length \(N\) to those of length about \(D/N\), where \(D\) measures the range in which the pairs \((s, \chi)\) may lie. It is useful to have a peak function which is itself a Dirichlet polynomial: we use the \(\mathcal{H}\) series discussed in Sections 3 and 4, which are modified Dirichlet \(L\)-functions. It is sometimes possible to use \(F(s, \chi)\) itself as a peak function, as in Lemma 10 below. The \(L\)-functions can be approximated by \(\mathcal{H}\) series of length \(D^{10}\) (the so-called approximate functional equation), as in Lemma 14 below. Lemma 14 is implicit in the literature; we sketch the proof out of duty. Jutila [9] has a new lemma (our Lemma 7) in which \(F(s, \chi)\) is raised to an even integral power, and obtains sharper results than those of \[8\] when \(F(s, \chi)\) is very large, for instance when the exponent \(10^{a}\) of \((2.23)\) is \(4/5\).

In this paper we explore the consequences of Jutila’s new lemma. Our arguments are purely combinatoric (except Lemma 14). To make the work accessible, we have summarised the main ideas of previous papers as a sequence of lemmas, stressing the combinatoric rather than the analytic aspects. Our result is Theorem 2 of Section 5. It enables us to improve the zero-density theorems for Dirichlet \(L\)-functions. For instance we extend the range of the density hypotheses. Let \(N(a, T, \chi)\) be the number of zeros \(\beta + iy\) of \(L(s, \chi)\) in \(\beta \geq a, |y| \leq T\). Then

\[
\sum_{\chi \mod q} N(a, T, \chi) \ll (qT^{2-2a})^{2}
\]

holds for \(a > 4/5\). Let an asterisk denote a sum over proper characters. Then

\[
\sum_{q \leq Q} \sum_{\chi \mod q} N(a, T, \chi) \ll (Q^{2}T^{2-2a+2})^{a}
\]