MATHEMATICS

ON THE ALGEBRAIC APPROXIMATION OF FUNCTIONS. I

BY

JOHN COATES

(Communicated by Prof. J. POPKEN at the meeting of January 29, 1966)

The initial ideas of this theory of algebraic approximation come from HERMITE's famous paper "Sur la fonction exponentielle" (3), which started the modern theory of transcendental numbers. Hermite's basic principle was as follows. For a given system of distinct complex numbers $\omega_1, \omega_2, \ldots, \omega_m$ and of non-negative integers $\varrho_1, \varrho_2, \ldots, \varrho_m$, with sum σ, he constructed a system of m polynomials

$$\mathfrak{a}_1(z),\ \mathfrak{a}_2(z),\ \ldots,\ \mathfrak{a}_m(z)$$

of degrees equal to $\sigma-\varrho_1, \sigma-\varrho_2, \ldots, \sigma-\varrho_m$, respectively, such that all the remainder functions

$$\mathfrak{w}_{jk}(z)=\mathfrak{a}_k(z)\,e^{\omega_j z}-\mathfrak{a}_j(z)\,e^{\omega_k z} \qquad (j,k=1,2,\ldots,m)$$

vanish at $z=0$ at least to the order $\sigma+1$. He then used this system of polynomials to study the arithmetic approximation of the exponential function. In a later paper [4], HERMITE introduced a second type of algebraic approximation of the exponential function, by constructing a second system of m polynomials

$$a_1(z),\ a_2(z),\ \ldots,\ a_m(z)$$

of degrees equal to $\varrho_1-1, \varrho_2-1, \ldots, \varrho_m-1$, respectively, such that the remainder function

$$r(z)=\sum_{k=1}^{m} a_k(z)\,e^{\omega_k z}$$

vanishes at $z=0$ at least to the order $\sigma-1$. Hermite did not use his second polynomial system to deduce arithmetic properties of the exponential function, nor did he seemingly realize that his two types of approximation were related.

Some fifty years later, Mahler, in his paper "Zur Approximation der Exponentialfunktion und des Logarithm", returned to Hermite's two types of algebraic approximation of the exponential function. He deduced further properties of the two polynomial systems and showed that they were fundamentally related. Using both of Hermite's polynomial systems, Mahler was able to obtain very strong arithmetic properties of the exponential and logarithmic functions.

Hermite's two polynomial systems were based on special identities connected with the exponential function, and many mathematicians thought that the theory was too specialized to be of general applicability. SIEGEL's [11] general method of 1929 in the theory of transcendental numbers was based on a different principle. However, at a course given at the University of Groningen in 1934–5, MAHLER generalized Hermite's two approximation systems as follows. He considered an arbitrary system of m functions

$$f_1(z), f_2(z), \ldots, f_m(z)$$

which are analytic in some domain of the complex domain. Given an arbitrary sequence of equal or distinct points

$$z_1, z_2, z_3, \ldots$$

in this domain, and non-negative integers $\varrho_1, \varrho_2, \ldots, \varrho_m$, with sum σ, there exists then a non-trivial system of polynomials

$$\mathfrak{a}_1(z), \mathfrak{a}_2(z), \ldots, \mathfrak{a}_m(z),$$

of degrees at most $\sigma-\varrho_1, \sigma-\varrho_2, \ldots, \sigma-\varrho_m$, respectively, such that all the remainder functions

$$\mathfrak{w}_{jk}(z) = \mathfrak{a}_j(z) f_k(z) - \mathfrak{a}_k(z) f_j(z) \qquad (j, k = 1, 2, \ldots, m)$$

vanish at least at all the points $z_1, z_2, \ldots, z_{\sigma+1}$, and also there exists a non-trivial system of polynomials

$$a_1(z), a_2(z), \ldots, a_m(z)$$

of degrees at most $\varrho_1-1, \varrho_2-1, \ldots, \varrho_m-1$, respectively, such that the remainder function

$$r(z) = \sum_{k=1}^{m} a_k(z) f_k(z)$$

vanishes at least at all the points $z_1, z_2, \ldots, z_{\sigma-1}$. Mahler now introduced the notion of a *perfect* system (for his definition see part VII of this paper) of analytic functions as a natural generalization of Hermite's results. The significant change was that he defined this notion of a perfect system in terms of the properties of the algebraic approximation alone. Mahler did not publish his theory. In his unpublished manuscript, he obtained the results of part VII of this paper for perfect systems of analytic functions. For such systems he also introduced the transformation matrices of part VI of this paper, and obtained a number of explicit expressions for these matrices.

In the present paper, I generalize Mahler's theory as follows. An axiomatic theory is given, which, in particular, holds for rings of analytic functions defined on certain subsets of fields, which are complete under a valuation. Under some circumstances, the theory also holds for rings of formal power series with coefficients in any field. This theory therefore contains the work of MAHLER and the later work of JAGER [2] as particular cases. A *local* property is defined to be a property at *one* system $\varrho_1, \varrho_2, \ldots, \varrho_m$, while a *global* property is a property at *infinitely* many systems $\varrho_1, \varrho_2, \ldots, \varrho_m$. With this terminology, Mahler's global notion of *perfectness* is generalized to the local notion of *normality*, while still retaining locally the essential properties of the approximation. The methods used are very simple, and are those introduced by Mahler in his unpublished manuscript. Mahler placed his unpublished manuscript at my disposal, and gave his free permission for the use of its results. For this kindness, and his generous advice, I wish to express my sincere thanks to professor Mahler.

A central problem of this theory is to determine at what systems of non-negative integers $\varrho_1, \varrho_2, \ldots, \varrho_m$, a given system of functions $f_1, f_2, \ldots, f_m$ is normal. As a contribution to this problem, I prove that the local notion of normality always implies certain global properties (see the Normality Zigzag Theorem in part V). I also prove new results on the exponential function.

The study of this algebraic approximation is interesting for the following reasons. Firstly, the approximation has considerable interest in itself. Even for the exponential function, our knowledge of the behaviour of the approximation is quite limited. This approximation gives generalizations of several important functions of classical analysis, for example, the Gamma functions and the Beta function. JAGER [2] has shown that a particular case of the approximation generalizes many classical results on the padé Table.

Secondly, the approximation is a powerful tool for studying questions of arithmetic approximation in the theory of transcendental numbers and the Thue-Siegel Theorem (see [5], [6], [7], [9]).

Thirdly, the methods used should be applicable to the study of other types of algebraic approximation, as, for example, the approximation of p-adic integers by rational integers.

I.

1. We begin by introducing two rings in which we shall be studying algebraic approximation. A non-archimedean valuation is defined on one ring, and a sequence of pseudo-valuations on the other, and these are shown to be related in a simple manner. In the next part, we shall define the algebraic approximation to be studied in terms of this valuation and this sequence of pseudo-valuations.

Firstly, we introduce the ring from which the approximating elements will be chosen. Let ω be a Euclidean domain with non-archimedean valuation $\overline{|\ |}$. This means that for all $a, b \in \omega$,

(1) if $a \neq 0$, $\overline{|a|}$ is a non-negative integer, but $\overline{|0|} = -\infty$;

(2) $\overline{|a+b|} \leqslant \max\{\overline{|a|}, \overline{|b|}\}$;

(3) $\overline{|ab|} = \overline{|a|} + \overline{|b|}$;

(4) there exists $c, d \in \omega$ such that

$$a = bc + d$$

where either $d = 0$ or $\overline{|d|} < \overline{|b|}$.

Then, since ω is a principal ideal of itself, it has a unit element 1. Let F be the set of all those elements of ω which satisfy

$$\alpha = 0 \text{ or } \overline{|\alpha|} = 0.$$

Then F is a field. For it is obvious that F is a subring of ω, and if α is any non-zero element, of F, then, by (4), there exists $\beta \in \omega$ such that

$$\alpha\beta = 1.$$

This equation implies that $\overline{|\beta|} = 0$, and hence α has an inverse β in F. From now on F will be called the *constant field*, and the elements of F will be called *constants*.

Let p be an any prime element of ω. We associate with p the valuation ord_p defined by

$$\text{if } a \neq 0,\ ord_p(a) = n\overline{|p|},\ ord_p(0) = \infty,$$

where n is the largest non-negative integer such that $p^{-n}a$ is still an element of ω. Since every non-zero element $a \in \omega$ has a unique factorization into a finite number of primes, at most finitely many of the values

$$ord_p(a)$$

are distinct from zero. Further, these values are linked by the fundamental equation

$$\sum_p ord_p(a) = \overline{|a|}. \tag{5}$$

The valuations ord_p will be of subordinate importance in this paper, and will only be used in the proof of the fundamental lemma later in this part.

2. Next let

$$\Pi\colon p_1, p_2, p_3, \dots$$

be an arbitrary infinite sequence of equal or distinct elements of ω, satisfying

$$\overline{|p_\lambda|} = 1 \qquad (\lambda = 1, 2, \dots).$$

Then every element of this sequence is a prime of ω. Put

$$\psi_0 = 1,\ \psi_\lambda = \prod_{l=1}^{\lambda} p_l \qquad (\lambda = 1, 2, \dots),$$

so that

$$\overline{|\lambda_\lambda|} = \lambda \qquad (\lambda = 0, 1, 2, \dots).$$

The algebraic approximation to be studied will be defined in terms of this sequence of primes Π. We now introduce the ring in which we shall study this approximation. We define Ω to be an integral domain which contains ω as a subring, and all of whose elements satisfy the following expansion condition in terms of the sequence of primes Π:

(6) *for every $f \in \Omega$ and for every prime $p_\lambda \in \Pi$, there exists a unique constant φ_λ and an element $f_\lambda \in \Omega$, both depending on f and p_λ, such that*

$$f = \varphi_\lambda + p_\lambda f_\lambda.$$

Before studying the properties of Ω, we note that, given ω, at least one such ring always exists, since the ring ω itself always satisfies this condition. However, from the point of view of algebraic approximation, the most interesting czse is when ω is a *proper* subring of Ω. It is essential to our theory that the expansion constant φ_λ in [6] is unique. We shall call Ω a *weld ring* of the sequence of primes Π.

We now investigate the properties of Ω. If f is any element of Ω, then there exists a unique constant $\varphi^{(0)}$ and an element $f^{(1)} \in \Omega$ such that

$$f = \varphi^{(0)} + p_1 f^{(1)}.$$

Since Ω is an integral domain, $f^{(1)}$ is also uniquely determined. Again, since $f^{(1)} \in \Omega$, there exist a unique constant $\varphi^{(1)}$ and a unique element $f^{(2)} \in \Omega$ such that

$$f^{(1)} = \varphi^{(1)} + p_2 f^{(2)}.$$

Thus there exist unique constants $\varphi^{(0)}$, $\varphi^{(1)}$ and a unique element $f^{(2)} \in \Omega$ such that

$$f = \sum_{\lambda=0}^{1} \varphi^{(\lambda)} \psi_\lambda = \psi_2 f^{(2)}.$$

A simple inductive argument shows that, for all positive integers n, there exist unique constants $\varphi^{(0)}, \varphi^{(1)}, \ldots, \varphi^{(n-1)}$ and a unique element $f^{(n)} \in \Omega$ such that

$$f = \sum_{\lambda=0}^{n-1} \varphi^{(\lambda)} \psi_\lambda + \psi_n f^{(n)}.$$

This expression will be called the *interpolation series* for f. The unique constants $\varphi^{(0)}, \varphi^{(1)}, \ldots, \varphi^{(n-1)}$ will be called the *coefficients* of f. The elements of Ω will henceforth be called *functions*.

The existence of this *unique* interpolation series for every function is the fundamental property of the weld ring Ω. As a first consequence of this property, we show that ω can be characterized as the set of all functions which have *finite* interpolation series, in the sense that only a finite number of their coefficients are non-zero. For, if a is any element of ω, it follows from § 1 (4) that there exist a constant $\alpha^{(0)}$ an element $a^{(1)} \in \omega$ such that

$$a = \alpha^{(0)} + p_1 a^{(1)} \text{ and } \overline{|a^{(1)}|} \leqslant \overline{|a|} - 1.$$

Repeating the above argument for $a^{(1)} \in \omega$, and so on, after a finite number of steps, we obtain the finite interpolation series

$$a = \sum_{\lambda=0}^{\overline{|a|}} \alpha^{(\lambda)} \psi_\lambda .$$

This is then the only interpolation series for a in the weld ring Ω. Since the valuation $\overline{|\ \ |}$ is non-archimedean, this interpolation series has the further property that $\alpha^{(\overline{|a|})} \neq 0$. Conversely, every finite interpolation series

$$b = \sum_{\lambda=0}^{n} \beta^{(\lambda)} \varphi_\lambda$$

is an element of ω, satisfying $\overline{|b|} \leqslant n$, with equality if and only if $\beta^{(n)} \neq 0$. For this reason, call the elements of ω *polynomials*, and for all polynomials a the value $\overline{|a|}$ will be called the *degree* of a.

Hence, in particular, for all non-negative integers n and m,

$$\psi_n \psi_m = \sum_{\lambda=\max\{m,n\}}^{m+n} \tau^{(\lambda)}(m, n)\, \psi_\lambda ,$$

where the $\tau^{(\lambda)}(m, n)$ are certain constants such that always $\tau^{(m+n)}(m, n)$ is non-zero. From now on we shall suppose that these constants are normalized so that

$$\tau^{(m+n)}(m, n) = 1 \qquad (m, n = 0, 1).$$

The interpolation series for the sum of two functions is equal to the sum of their respective interpolation series. Similarly, using the expressions for the interpolation series of the products $\psi_n \psi_m$, we can oblain the interpolation series for the product of two functions from their respective interpolation series.

3. Let θ be either Ω or ω, and let f by any element of θ. Then the symbol $(f)_\theta$ will denote the principal ideal of θ generated by f. The symbol (f) will only be used to denote the principal ideal of Ω generated by f.

We now define a sequence of pseudo-valuations of Ω

$$\underline{|\ \ \ |}_n \qquad (n = 1, 2, \ldots)$$

as follows. For every $f \in \Omega$ and for every positive integer n, put

$$\underline{|f|}_n = \lambda \text{ if } f \in (\psi_\lambda^n), \text{ but } f \notin (\psi_{\lambda+1}^n).$$

Then each of the pseudo-valuations $\underline{|\ \ |}_n$ satisfies, for all $f, g \in \Omega$,

(7) $$\underline{|f|}_n \geqslant 0,\ \underline{|0|}_n = \infty;$$

(8) $$\underline{|f+g|}_n \geqslant \min\{\underline{|f|}_n, \underline{|g|}_n\};$$

(9) $$\underline{|fg|}_n \geqslant \max\{\underline{|f|}_n, \underline{|g|}_n\}.$$

The first pseudo-valuation $\underline{|\ \ |}_1$ has a particularly important role in the subsequent theory. For brevity, we shall often write $\underline{|\ \ |}$ instead of $\underline{|\ \ |}_1$, and we shall call the value $\underline{|f|}$ the *order* of f.

The valuation $\overline{|\ \ |}$ and each of the pseudo-valuations $\underline{|\ \ |}_n$ are closely related, and this relation is the basis of most of the investigations of this paper. Explicitly, the relation is the following.

Lemma 1. *If λ is any non-negative integer, and a is a polynomial satisfying*

$$\overline{|a|} \leqslant n\lambda, \ \underline{|a|}_n \geqslant \lambda,$$

then $a = \alpha\psi_\lambda^n$, for some constant α. Further, α is non-zero if and only if $\overline{|a|} = n\lambda$.

Proof. Firstly, we use induction on n to prove that for all positive integers n

$$(\psi_\lambda^n)_\Omega \cap \omega = (\psi_\lambda^n)_\omega \qquad (\lambda = 0, 1, \ldots).$$

It is obvious that for all positive integers n

$$(\psi_\lambda^n)_\omega \subseteq (\psi_\lambda^n)_\Omega \cap \omega \qquad (\lambda = 0, 1, \ldots).$$

We prove the converse by induction on n. Consider first the case $n = 1$. If b is any element of $(\psi_\lambda)_\Omega \cap \omega$, then the interpolation series for b must be of the form

$$b = \sum_{l=\lambda}^{\overline{|b|}} \beta^{(l)} \psi_l.$$

It is therefore immediate that $b \in (\psi_\lambda)_\omega$, and this $(\psi_\lambda)_\Omega \cap \omega \subseteq (\psi_\lambda)_\omega$. This proves the case $n = 1$. Next assume the result true for $n-1$. If b is any element of $(\psi_\lambda^n)_\Omega \cap \omega$, then, in particular, b is an element of $(\psi_\lambda)_\Omega \cap \omega$, so that b is an element of the ideal $(\psi_\lambda)_\omega$. Therefore b/ψ_λ is a polynomial, and since Ω is an integral domain, $b/\psi_\lambda \in (\psi_\lambda^{n-1})_\Omega \cap \omega$. Thus, by the inductive hypothesis, $b/\psi_\lambda \in (\psi_\lambda^{n-1})_\omega$, whence $b \in (\psi_\lambda^n)_\omega$. We have therefore shown that

$$(\psi_\lambda^n)_\Omega \cap \omega \subseteq (\psi_\lambda^n)_\omega \qquad (\lambda = 0, 1, \ldots),$$

and this completes the proof of the first assertion.

Secondly, if a is a polynomial satisfying the hypotheses then, by the definition of the pseudo-valuation $\underline{|\ \ |}_n$, $a \in (\psi_\lambda^n)_\Omega \cap \omega$. Since $(\psi_\lambda^n)_\Omega \cap \omega = (\psi_\lambda^n)_\omega$, it follows that $a = b\psi_\lambda^n$, where b is a polynomial. We can suppose that a is non-zero, because the lemma holds for $a = 0$. The fundamental equation (5) therefore gives

$$n\lambda + \sum_p ord_p(b) = \overline{|a|} \leqslant n\lambda,$$

and this inequality implies that

$$\overline{|a|} = n\lambda, \ \overline{|b|} = \sum_p ord_p(b) = 0.$$

This completes the proof.

II.

4. We now establish two distinct types of simultaneous algebraic approximation of functions by polynomials. These two types of approximation are fundamentally related, as will be shown later.

From now on, m will denote a fixed, but arbitrary, positive integer such that $m \geqslant 2$, because the case $m=1$ is trivial.

Let

$$f_1, f_2, \ldots, f_m$$

be any fixed system of m functions. For brevity, we shall henceforth call this system of functions a *function vector*, and denote the whole system by the symbol $\boldsymbol{f}$. For all positive integers q, the function vector

$$f_1^q, f_2^q, \ldots, f_m^q$$

will be denoted by the symbol $\boldsymbol{f}^q$.

Let further

$$\varrho_1, \varrho_2, \ldots, \varrho_m$$

be a system of m parameters, with sum σ, which will always be assumed to range over the non-negative integers. This system of m parameters will be said to be *trivial* if

$$\varrho_1 = \varrho_2 = \ldots = \varrho_m = 0.$$

Likewise, a system of m polynomials will be said to be *trivial* if all of its elements are zero.

With this notation, there exists then a non-trivial system of polynomials

$$a_k(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (k=1, 2, \ldots, m),$$

which, together with the remainder function

$$r(\varrho_1 \varrho_2 \cdots \varrho_m) = \sum_{k=1}^{m} a_k(\varrho_1 \varrho_2 \cdots \varrho_m) f_k,$$

satisfies the inequalities

$$\overline{|a_k(\varrho_1 \varrho_2 \cdots \varrho_m)|} \leqslant \varrho_k - 1 \qquad (k=1, 2, \ldots, m),$$

$$\underline{|\ r(\varrho_1 \varrho_2 \cdots \varrho_m)|} \geqslant \sigma - 1.$$

For take m polynomials, with unknown coefficients, satisfying the former condition. These polynomials have $\sum_{k=1}^{m} \varrho_k = \sigma$ unknown coefficients amongst them. The function $r(\varrho_1 \varrho_2 \ldots \varrho_m)$ will satisfy the latter condition if and only if the coefficients of

$$\psi_0, \psi_1, \cdots \psi_{\sigma-2}$$

in its interpolation series are all zero. This gives $\sigma - 1$ homogeneous linear equations for these σ unknown coefficients. These equations always have a non-trivial solution in the field. F Hence the assertion is true. This is the first type of algebraic approximation, and we introduce the following terminology. The polynomial system and the remainder function

$$a_k(\varrho_1 \varrho_2 \cdots \varrho_m),\ r(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (k=1, 2, \ldots, m)$$

will be called a *Latin polynomial system* and its *remainder*, respectively, belonging to the system

$$\varrho_1, \varrho_2, \ldots, \varrho_m.$$

Further, there exists a non-trivial system of polynomials

$$\mathfrak{a}_k(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (k=1, 2, \ldots, m),$$

which, together with the remainder functions

$$\mathfrak{w}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) = \mathfrak{a}_k(\varrho_1 \varrho_2 \cdots \varrho_m)\, f_j - \mathfrak{a}_j(\varrho_1 \varrho_2 \cdots \varrho_m)\, f_k \qquad (j, k=1, 2, \ldots, m),$$

satisfies the inequalities

$$\overline{|\ \mathfrak{a}_k(\varrho_1 \varrho_2 \cdots \varrho_m)|} \leqslant \sigma - \varrho_k \qquad (k=1, 2, \ldots, m),$$

$$\underline{|\mathfrak{w}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m)|} \geqslant \sigma + 1 \qquad (j, k=1, 2, \ldots, m).$$

For take m polynomials with unknown coefficients, satisfying the former condition. These polynomials have $\sum_{k=1}^{m} \sigma - \varrho_k + 1 = (m-1)(\sigma+1)+1$ unknown coefficients amongst them. The remainder functions

$$\mathfrak{w}_{jk}(\varrho_1 \varrho_2, \ldots, m) \qquad (j, k=1, 2, \ldots, m)$$

satisfies the latter condition if and only if the coefficients of

$$\psi_0, \psi_1, \ldots, \psi_\sigma$$

in the interpolation series for

$$\mathfrak{w}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (j=1, 2, \ldots, m, j \neq k)$$

are all zero. This gives $(m-1)(\sigma+1)$ homogeneous linear equations for these $(m-1)(\sigma+1)+1$ unknown coefficients, which always have a non-trivial solution in the field F. Hence the assertion is true. Again, we introduce the following terminology. The polynomial system and the remainder functions

$$\mathfrak{a}_k(\varrho_1 \varrho_2 \cdots \varrho_m),\ \mathfrak{w}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (j, k=1, 2, \ldots, m)$$

will be called a *German polynomial system* and its *remainders*, respectively, belonging to the system

$$\varrho_1, \varrho_2, \ldots, \varrho_m$$

5. To study the properties of the approximation just introduced, it is now necessary to impose a minor restriction on the function vector $\boldsymbol{f}$.

We shall say that the function vector $\boldsymbol{f}$ *vanishes at the prime* $p_\lambda \in \Pi$ if the m expansion constants $\varphi_{1\lambda}, \varphi_{2\lambda}, \ldots, \varphi_{m\lambda}$, where

$$f_k = \varphi_{k\lambda} + p_\lambda f_{k\lambda} \qquad (k=1, 2, \ldots, m)$$

are all zero. The result we need is the following.

Lemma 2. *If the function vector* $\boldsymbol{f}$ *vanishes at none of the primes in* Π, *then, for all positive integers* n *and* q, *and for all functions* g, *the inequalities*

$$\underline{|gf_j^q|}_n \geqslant \lambda \qquad (j=1, 2, \ldots, m)$$

imply

$$\underline{|g|}_n \geqslant \lambda.$$

Proof. The proof of the lemma is in three parts. Firstly, we assume that $n=1$ and $q=1$, and use induction on λ. The result is trivial correct when $\lambda=0$, and thus we suppose it true for $\lambda-1$. Let g be any function satisfying $\underline{|gf_j|} \geqslant \lambda$, or equivalently $gf_j \in (\psi_\lambda)$, for $j=1, 2, \ldots, m$. Then, by the inductive hypothesis, $g \in (\psi_{\lambda-1})$. In addition, there exist the expansions

$$f_j = \varphi_{j\lambda} + p_\lambda f_{j\lambda} \qquad (j=1, 2, \ldots, m),$$

where, by assumption, at least one of the constants $\varphi_{j\lambda}$ is non-zero. It is therefore clear that $g \in (\psi_\lambda)$, which is the assertion for λ. This completes the proof of the first part.

Secondly, the result just proven implies the assertion for all positive integers q, since, if the function vector $\boldsymbol{f}$ vanishes at none of the primes in Π, then the function vector $\boldsymbol{f}^q$ likewise vanishes at none of the primes in Π. We are still assuming that $n=1$.

Finally, an obvious inductive argument (just as in Lemma 1) proves the assertion for arbitrary positive integers n. This completes the proof.

Henceforth, unless stated to the contrary, we sha.l always assume that the function vector $\boldsymbol{f}$ vanishes at none of the primes in Π.

6. We now introduce an expression which will be of fundamental importance in our study of the algebraic approximation. The expression

$$e\begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix}$$

will always denote an expression of the form

$$e\begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix} = \sum_{k=1}^{m} a_k \mathfrak{a}_k,$$

where

$$a_k \ , \ \mathfrak{a}_k \qquad (k=1, 2, \ldots, m)$$

are two systems of polynomials, which, together with the functions

$$r = \sum_{k=1}^{m} a_k f_k \ , \ \mathfrak{w}_{jk} = \mathfrak{a}_k f_j - \mathfrak{a}_j f_k \qquad (j, k = 1, 2, \ldots, m),$$

satisfy the inequalities

$$\overline{|a_k|} \leqslant r_k - 1 \ , \ \overline{|\ \mathfrak{a}_k\ |} \leqslant \mathfrak{s} - \mathfrak{w}_k,$$

$$\underline{|\ r\ |} \geqslant s - 1 \ , \ \underline{|\mathfrak{w}_{jk}|} \geqslant \mathfrak{s} + 1.$$

From these inequalities, we at once obtain the following estimate,

$$(D) \qquad \overline{\left| e \begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix} \right|} \leqslant \max_{k=1,2,m} \{(r_k - 1) + (\mathfrak{s} - \mathfrak{w}_k)\}.$$

Further, for $j = 1, 2, \ldots, m$,

$$f_j\, e \begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix} = \mathfrak{a}_j r + \sum_{k=1}^{m} \mathfrak{w}_{jk} a_k,$$

and from it we deduce that

$$\underline{\left| f_j\, e \begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix} \right|} \geqslant \min\{s - 1, \mathfrak{s} + 1\} \qquad (j = 1, 2, \ldots, m).$$

But the function vector $\boldsymbol{f}$ vanishes at none of the primes in Π, and thus, by Lemma 2,

$$(0) \qquad \underline{\left| e \begin{pmatrix} r_1 r_2 \ldots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \ldots \mathfrak{w}_m \mathfrak{s} \end{pmatrix} \right|} \geqslant \min\{s - 1, \mathfrak{s} + 1\}.$$

The estimates (D) and (0) are basic for our later work.

III.

7. We begin with two definitions. A *local* property of the approximation is defined to be a property at *one* system $\varrho_1, \varrho_2, \ldots, \varrho_m$, while a *global* property is a property which holds for *infinitely* many systems $\varrho_1, \varrho_2, \ldots, \varrho_m$. We derive three global properties of the approximation which hold for every function vector that vanishes at none of the primes in Π.

For a given system $\varrho_1, \varrho_2, \ldots, \varrho_m$ let us arbitrarily choose one (there may be many) system of Latin polynomials, which is then supposed fixed for the rest of this part. We denote this system and its remainder by

$$a_k(\varrho_1 \varrho_1 \cdots \varrho_m) \ , \ r(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (k = 1, 2, \ldots, m).$$

Let $\delta_{hk} = \begin{cases} 1 \text{ if } h = k \\ 0 \text{ if } h \neq k \end{cases}$ be the Kronecker symbol. For $h = 1, 2, \ldots, m$ put

$$a_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m) = a_k(\varrho_1 + \delta_{h1}\, \varrho_2 + \delta_{h2} \cdots \varrho_m + \delta_{hm}) \qquad (k = 1, 2, \ldots, m),$$

$$r_h(\varrho_1 \varrho_2 \cdots \varrho_m) = r(\varrho_1 + \delta_{h1}\, \varrho_2 + \delta_{h2} \cdots \varrho_m + \delta_{hm}),$$

so that

$$r_h(\varrho_1 \varrho_2 \cdots \varrho_m) = \sum_{k=1}^{m} a_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m) f_k.$$

Further, let $a(\varrho_1\, \varrho_2 \,\ldots\, \varrho_m)$ be the $m \times m$ matrix

$$a(\varrho_1 \varrho_2 \cdots \varrho_m) = (a_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m))_{h,k=1,2,\ldots,m},$$

and let $d(\varrho_1\, \varrho_2 \,\ldots\, \varrho_m)$ be the determinant of this matrix.

This determinant can easily be evaluated. For the degree of the element in its h^{th} row and kth column is at most $\varrho_k + \delta_{hk} - 1$. Thus the degree of the determinant does not exceed the sum of the greatest possible degrees of its diagonal elements, that is

$$\overline{|d(\varrho_1 \varrho_2 \cdots \varrho_m)|} \leqslant \sigma.$$

Next we obtain an estimate for the order of the determinant. Applying elementary column operations to the determinant, we find that for $j = 1, 2, \ldots, m$.

$$f_j\, d(\varrho_1 \varrho_2 \cdots \varrho_m) = \begin{vmatrix} a_{11}, & \ldots, & a_{1j-1}, & r_1, & a_{1j+1}, & \ldots, & a_{1m} \\ a_{21}, & \ldots, & a_{2j-1}, & r_2, & a_{2j+1}, & \ldots, & a_{2m} \\ & & & & & & \\ a_{m1}, & \ldots, & a_{mj-1}, & r_m, & a_{mj+1}, & \ldots, & a_{mm} \end{vmatrix},$$

where the parameters have been omitted for brevity. Now the remainders satisfy

$$\underline{|r_h(\varrho_1 \varrho_2 \cdots \varrho\ \ )|} \geqslant \sigma \qquad (h = 1, 2, \ldots, m),$$

so that

$$\underline{|f_j\, d(\varrho_1 \varrho_2 \cdots \varrho_m)|} \geqslant \sigma \qquad (j = 1, 2, \ldots, m).$$

By assumption, the function vector $\boldsymbol{f}$ vanishes at none of the primes in Π, whence we obtain the estimate

$$\underline{|d(\varrho_1 \varrho_2 \cdots \varrho_m)|} \geqslant \sigma.$$

Therefore

$$d(\varrho_1 \varrho_2 \cdots \varrho_m) = \alpha \psi_\sigma \text{ with } \alpha \in F.$$

The constant α will be non-zero if and only if the degree of the determinant is equal σ, and this will be so if and only if

$$\overline{|a_{hh}(\varrho_1 \varrho_2 \cdots \varrho_m)|} = \varrho_h \qquad (h = 1, 2, \ldots, m).$$

We shall investigate this property and its implications later.

8. There is an analogous determinant in the German polynomials and we can evaluate it in a similar manner. For a given system $\varrho_1, \varrho_2, \ldots, \varrho_m$ let us arbitrarily choose one system of German polynomials, which are then again supposed fixed for the rest of this part. We denote this system

and its remainders by

$$\mathfrak{a}_k(\varrho_1 \varrho_2 \cdots \varrho_m) \; , \; \mathfrak{w}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (j, k = 1, 2, \ldots, m).$$

For $h = 1, 2, \ldots, m$, put

$$\mathfrak{a}_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m) = \mathfrak{a}_k(\varrho_1 - \delta_{h1}\, \varrho_2 - \delta_{h2} \cdots \varrho_m - \delta_{hm}) \qquad (k = 1, 2, \ldots, m),$$

$$\mathfrak{w}_{hjk}(\varrho_1 \varrho_2 \cdots \varrho_m) = \mathfrak{w}_{jk}(\varrho_1 - \delta_{h1}\, \varrho_2 - \delta_{h2} \cdots \varrho_m - \delta_{hm}) \qquad (j, k = 1, 2, \ldots, m),$$

so that

$$\mathfrak{w}_{hjk}(\varrho_1 \varrho_2 \cdots \varrho_m) = \mathfrak{w}_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m)\, f_j - \mathfrak{a}_{hj}(\varrho_1 \varrho_2 \cdots \varrho_m)\, f_k \qquad (j, k = 1, 2, \ldots, m).$$

Further, let $\mathfrak{a}(\varrho_1, \varrho_2 \ldots \varrho_m)$ be the $m \times m$ matrix

$$\mathfrak{a}(\varrho_1 \varrho_2 \cdots \varrho_m) = (\mathfrak{a}_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m))_{h,k=1,2,\ldots,m},$$

and let $\mathfrak{d}(\varrho_1\, \varrho_2 \ldots \varrho_m)$ be the determinant of this matrix.

In this determinant, the degree of the element in the h^{th} row and k^{th} column is at most $\sigma - 1 - \varrho_k + \delta_{hk}$. Therefore the degree of the determinant does not exeed the sum of the greatest possible degrees of its diagonal elements, that is again

$$\overline{|\mathfrak{d}(\varrho_1 \varrho_2 \cdots \varrho_m)|} \leqslant (m-1)\,\sigma.$$

Next, applying again elementary column operations to the determinant, we find that for $j = 1, 2, \ldots, m$

$$f_j^{m-1}\,\mathfrak{d}(\varrho_1 \varrho_2 \cdots \varrho_m) = \begin{vmatrix} \mathfrak{w}_{1j1}, & \ldots, & \mathfrak{w}_{1jj-1}, & \mathfrak{a}_{1j}, & \mathfrak{w}_{1jj+1}, & \ldots, & \mathfrak{w}_{1jm} \\ \mathfrak{w}_{2j1}, & \ldots, & \mathfrak{w}_{2jj-1}, & \mathfrak{a}_{2j}, & \mathfrak{w}_{2jj+1}, & \ldots, & \mathfrak{w}_{2jm} \\ & & & & & & \\ \mathfrak{w}_{mj1}, & \ldots, & \mathfrak{w}_{mjj-1}, & \mathfrak{a}_{mj}, & \mathfrak{w}_{mjj+1}, & \ldots, & \mathfrak{w}_{mjm} \end{vmatrix},$$

where the parameters have been omitted for brevity. Now the remainders satisfy

$$\underline{|\mathfrak{w}_{hjk}(\varrho_1 \varrho_2 \cdots \varrho_m)|} \geqslant \sigma \qquad (h, j, k = 1, 2, \ldots, m),$$

so that

$$\underline{|f_j^{m-1}\,\mathfrak{d}(\varrho_1 \varrho_2 \cdots \varrho_m)|}_{m-1} \geqslant \sigma \qquad (j = 1, 2, \ldots, m).$$

By assumption, the function vector f vanishes at none of the primes in Π, whence, by Lemma 2, we obtain the estimate

$$\underline{|\mathfrak{d}(\varrho_1 \varrho_2 \cdots \varrho_m)|}_{m-1} \geqslant \sigma.$$

Therefore the value of the determinant is

$$\mathfrak{d}(\varrho_1 \varrho_2 \cdots \varrho_m) = \beta \psi_\sigma^{m-1}, \text{ with } \beta \in F.$$

The constant β will be non-zero if and only if the degree of the determinant is equal to $(m-1)\sigma$, and this will be so if and only if

$$\overline{|\mathfrak{a}_{hh}(\varrho_1\, \varrho_2 \cdots \varrho_m)|} = \sigma - \varrho_h \qquad (h = 1, 2, \ldots, m).$$

We shall investigate also this property and its implications later.

8. Finally, we prove a remarkable relation between the Latin and German polynomial systems. The polynomials

$$e_{hj} = \sum_{k=1}^{m} a_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m)\, \mathfrak{a}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) \qquad (h,j=1,2,\dots,m)$$

are expressions of the form $e\begin{pmatrix} r_1 r_2 \dots r_m s \\ \mathfrak{w}_1 \mathfrak{w}_2 \dots \mathfrak{w}_m \mathfrak{s} \end{pmatrix}$, with parameter values

$$r_k = \varrho_k + \delta_{hk} \;,\; \mathfrak{w}_k = \varrho_k - \delta_{jk},$$

$$s = \sigma + 1 \quad , \quad \mathfrak{s} = \sigma - 1.$$

From these values, (D) and (0) give the estimates

$$\overline{|e_{hj}|} \leqslant \max_{k=1,\dots,m} \{(\varrho_k + \delta_{hk} - 1) + (\sigma - 1 - \varrho_k + \delta_{jk})\} = \max_{k=1,\dots,m} \{\sigma + \delta_{hk} + \delta_{jk} - 2\},$$

$$\underline{|e_{hj}|} \geqslant \min\{\sigma, \sigma\} = \sigma.$$

These estimates imply that for $h \neq j$

$$\overline{|e_{hj}|} < \underline{|e_{hj}|} \;,\; \text{i.e. } e_{hj} = 0,$$

and for $h = j$

$$\sigma \geqslant \overline{|e_{hh}|} \leqslant \underline{|e_{hh}|} \geqslant \sigma \;,\; \text{i.e. } e_{hh} = \varepsilon_h \psi_\varrho, \text{ with } \varepsilon_h \in F.$$

The constant ε_h is non-zero if and only if the degree of e_{hk} is equal to σ, and this is so if and only if

$$\overline{|a_{hh}(\varrho_1 \varrho_2 \cdots \varrho_m)|} = \varrho_h \;,\; \overline{|\mathfrak{a}_{hh}(\varrho_1 \varrho_2 \cdots \varrho_m)|} = \sigma - \varrho_h.$$

We shall investigate also this property later. In any case,

$$\sum_{k=1}^{m} a_{hk}(\varrho_1 \varrho_2 \cdots \varrho_m)\, \mathfrak{a}_{jk}(\varrho_1 \varrho_2 \cdots \varrho_m) = \delta_{hj}\, \varepsilon_h\, \psi_\sigma \qquad (h,j=1,2,\dots,m),$$

and these equations can be written as the single matrix equation

$$a(\varrho_1 \varrho_2 \cdots \varrho_m)\, \mathfrak{a}'(\varrho_1 \varrho_2 \cdots \varrho_m) = \psi_\sigma \begin{pmatrix} \varepsilon_1, & 0, & \dots, & 0 \\ 0, & \varepsilon_1, & \dots, & 0 \\ 0, & 0, & \dots, & \varepsilon_m \end{pmatrix}$$

where the dash denotes the transposed matrix.

(To be continued)