*J. Austral. Math. Soc. (Series A)* **51** (1991), 50–61

# GENERALISED SIMULTANEOUS APPROXIMATION OF FUNCTIONS

A. J. VAN DER POORTEN

(Received 11 October 1989)

Communicated by J. H. Loxton

*Dedicated to the memory of Kurt Mahler*

**Abstract**

We generalise the approximation theory described in Mahler's paper "Perfect Systems" to linked simultaneous approximations and prove the existence of nonsingular approximation and of transfer matrices by generalising Coates' normality zig-zag theorem. The theory sketched here may have application to constructions important in the theory of diophantine approximation.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*) (1985 *Revision*): 11 J 61.

In 1934, at the University of Groningen, Kurt Mahler gave a course in which he generalised the approximation theory for the exponential function presented in Hermite's fundamental papers [2], [3]. It had seemed that Hermite's theory was unique to the exponential function, but Mahler showed that the theory is quite general and has broad application in the theory of diophantine approximation as well as providing a generalisation of the Padé theory now well known to applied mathematicians; *cf.* [5]. Mahler's lectures were not published until many years later [7]. However, in the meantime the manuscript provided a basis for the thesis of Henk Jager [4] and the honours essay (and immediately subsequent work) of John Coates [1].

The present note generalises remarks in [6], which themselves use the ideas of Coates [1] (particularly his notion of *normality*) as their starting point. The principal innovation of that joint paper [6] was that it provides an explicit

This research was partially supported by the Australian Research Council.

application of the formal inexplicit general theory to a problem in transcendence theory.

We do not consider the axiomatic theory provided by Coates [1]; however, our relatively formal treatment readily allows its restoration. Our purpose is to develop the generalised theory to a stage that demonstrates the existence of transfer matrices that, in principle, permit a sequential construction of the approximating polynomials. The author recalls conversations with Kurt Mahler in which Mahler expressed confidence in those notions and regret at the neglect of that aspect of his investigations.

The motivation for the present generalisation of Mahler's theory is the problem of constructing a polynomial in several variables with prescribed vanishing. As shown in the concluding section, that construction can lead to a linked approximation problem of the kind about to be described. The central result below, *the normality zig-zag theorem* shows that, conveniently, the approximation systems often are unique.

## 1. Formal series

Let $\mathbf{F}$ be a field and, for $j = 0, 1, \dots, n$, let $w_j = (w_{1j}, w_{2j}, \dots)$ be sequences of points of $\mathbf{F}$; write $w = (w_0, \dots, w_n)$. We say that $f$ is a formal series over $\mathbf{F}$ with respect to $w$ if $f$ belongs to some ring containing $\mathbf{F}[z]$ such that, for $j = 0, 1, \dots, n$, and each integer $h \geq 1$, $f$ can be written uniquely as

$$f(z) = c_{0j} + c_{1j}(z - w_{1j}) + c_{2j}(z - w_{1j})(z - w_{2j}) \\ + \cdots + c_{h-1,j}(z - w_{1j}) \cdots (z - w_{h-1,j}) + (z - w_{1j}) \cdots (z - w_{hj}) g_{h,j},$$

where $g_{h,j}$ belongs to the same ring and the elements $c_{0j}, c_{1j}, \dots$ belong to $\mathbf{F}$. In the sequel $w$ is fixed. We write, for $j = 0, \dots, n$,

$$\operatorname{ord}_j f = \sigma_j$$

if $c_{0j} = a_{1j} = \cdots = c_{\sigma_j - 1, j} = 0$ and $c_{\sigma_j, j} \neq 0$.

In applications in the theory of diophantine approximation one seems, invariably, to take all the $w_{ij}$ equal to $0$. This motivates the terminology and notation used here and below. Nevertheless, the apparent generality of our setting is not spurious since it facilitates seeing various results that are otherwise more difficult to notice.

https://doi.org/10.1017/S1446788700033292 Published online by Cambridge University Press

## 2. Simultaneous approximation

Let $F$ be an $(n+1)\times(n+1)$ matrix of formal series $F=(f_{ij})$. We describe $F$ as *generic* if the series comprising each row of $F$ are linearly independent over the ring $\mathbf{F}[z]$ of polynomials with coefficients in $\mathbf{F}$ and if there is a a permutation $\pi$ of $\{0,\dots,n\}$ so that

$$f_{0,\pi(0)}(w_{k_0,0})f_{1,\pi(1)}(w_{k_1,1})\cdots f_{n,\pi(n)}(w_{k_n,n})$$

does not vanish for any selection of elements $w_{k_j,j}$ respectively from the sequences $w_j$.

One says that a vector $a=(a_0,\dots,a_n)$ of polynomials $a_i$ is an *approximation* to $F$ *at* $\rho=(\rho_0,\dots,\rho_n)$ of *order* $\sigma=(\sigma_0,\dots,\sigma_n)$ if for $j=0,1,\dots,n$,

$$\operatorname{ord}_j\left(\sum_{i=0}^n a_i f_{ij}\right)\geq \sigma_j-\delta_{0j}$$

and for $i=0,1,\dots,n$,

$$\deg a_i\leq \rho_i-1.$$

Existence is a matter of $\sum\rho_i$ free coefficients which are to satisfy $\sum(\sigma_j-\delta_{0j})=|\sigma|-1$ linear conditions. Set $|\rho|=\sum\rho_i$. Then $F$ has an approximation at $\rho$ of order $\sigma$ whenever $|\rho|\geq|\sigma|-1$. In the sequel we always take $|\sigma|=|\rho|$.

It will be convenient to write

$$\psi_j(z|\sigma_j)=(z_j-w_{1j})\cdots(z-w_{\sigma_j-\delta_{0j},j})$$

and

$$\Psi(z|\sigma)=\prod_{j=0}^n\psi_j(z|\sigma_j).$$

Then $\Psi(z|\sigma)$ is a monic polynomial in $z$ of degree $|\sigma|-1=|\rho|-1$.

## 3. Approximation matrices

In the sequel $J$ denotes some nonempty subset of $\{0,1,\dots,n\}$.

Suppose that $a$ and $a'$ are approximations to $F$ at $\rho$ of order $\sigma$; or, for brevity, at $(\rho,\sigma)$. We say that the approximation $a'$ lies $J$-above the approximation $a$ if, for each $j\in J$,

$$\operatorname{ord}_j\left(\sum_i a_i' f_{ij}\right)\geq \operatorname{ord}_j\left(\sum_i a_i f_{ij}\right).$$

https://doi.org/10.1017/S1446788700033292 Published online by Cambridge University Press

This definition induces a partial ordering on the approximations at $(\rho, \sigma)$. We say that $a$ is a $J$-best approximation to $F$ at $(\rho, \sigma)$ if it is a maximal element with respect to this ordering.

In the present spirit we also say that an $(n+1)$-tuple $\phi' = (\phi'_0, \dots, \phi'_n)$ lies *above* an $(n+1)$-tuple $\phi = (\phi_0, \dots, \phi_n)$ if $\phi'_j \geq \phi_j$ for all $j = 0, \dots, n$; $\phi'$ lies $J$-above $\phi$ if the inequalities hold at least for each $j \in J$. Note that this entails that $\phi$ lies $J$-above itself for all $J$.

Suppose that every approximation $a$ at $(\rho, \sigma)$ satisfies

$$\operatorname{ord}_j \left( \sum_i a_i f_{ij} \right) = \sigma_j - \delta_{0j}$$

for each $j \in J$. Then we say that $F$ has *$J$-normal* approximation at $(\rho, \sigma)$. Up to normalisation by multiplication by an element of $\mathbf{F}^\times$, the *$J$-normal* approximation $a$ at $(\rho, \sigma)$ is unique. To see that, note that some linear combination of essentially distinct approximations yields an approximation strictly $J$-above $a$.

Fix some $l \in J$ and for each $h = 0, \dots, n$ consider approximations $(a_{h0}, \dots, a_{hn})$ to $F$ with order $\sigma + \delta_l = (\sigma_0 + \delta_l, \dots, \sigma_n + \delta_l)$ at $\rho + \delta_h = (\rho_0 + \delta_{0h}, \dots, \rho_n + \delta_{nh})$. Because $|\rho| + 1 \geq |\sigma|$, the existence of such approximations is guaranteed. Moreover, for each $h$, we must have $\deg a_{hh} = \rho_h$ since otherwise there is an approximation at $(\rho, \sigma)$ strictly $J$-above the unique $J$-best approximation. Accordingly we may normalise the approximations by multiplication by appropriate elements of $\mathbf{F}^\times$. We denote the normalised approximations, with the coefficient of $z^{\rho_h}$ of $a_{hh}$ equal to 1, by $(A_{h0}, \dots, A_{hn})$.

We call the $(n+1) \times (n+1)$ matrix

$$A(z) = A(z|\rho, \sigma; l) = \big(A_{hi}(z|\rho, \sigma; l)\big)$$

the *approximation matrix* to $F$ at $(\rho, \sigma; l)$.

**Theorem** (Normality zig-zag theorem for matrices of formal series). *Let $F$ be a generic matrix of formal series. Then $F$ has nonsingular approximation matrices*

$$A\big(z|\rho(k), \sigma(k); l(k)\big)$$

*at infinite sequences of pairs*

$$\big(\rho(0), \sigma(0)\big) = (0, 0), \big(\rho(1), \sigma(1)\big), \big(\rho(2), \sigma(2)\big), \dots$$

*with $\rho(k+1)$ lying above $\rho(k)$, $\sigma(k+1)$ lying above $\sigma(k)$, and $|\rho(k)| = |\sigma(k)| = k$ for all $k$. Indeed, every nonsingular approximation matrix*

https://doi.org/10.1017/S1446788700033292 Published online by Cambridge University Press

*$A(z|\rho, \sigma; l)$ has $(\rho, \sigma)$ on such a generalised normality zig-zag. Moreover, there are generalised normality zig-zags with $\min_{0\le i\le n} \rho_i(k) \to \infty$ as $k \to \infty$. Furthermore, for each $k$ we have*

$$\det A\big(z|\rho(k), \sigma(k), l(k)\big) = \Psi(z|\sigma + \delta_{l(k)}).$$

Our proof is contained in the remarks below; they continue the discussion preceding the statement of the Theorem.

Consider the determinant $\det A(z)$. Each column has a unique entry $A_{hh}$ of maximal degree $\rho_h$, and that entry is monic. Hence

$$\det A(z) = z^{|\rho|} + \text{lower order terms in } z.$$

In particular, the determinant does not vanish. For $h, j = 0, \dots, n$, set

$$R_{hj} = \sum_i A_{hi} f_{ij}.$$

Then for any permutation $\pi$ of $\{0, \dots, n\}$ we obtain

$$f_{0,\pi(0)}(z) f_{1,\pi(1)}(z) \cdots f_{n,\pi(n)}(z) \det A(z) = \begin{vmatrix} R_{00} & R_{01} & \cdots & R_{0n} \\ R_{10} & R_{11} & \cdots & R_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n0} & R_{n1} & \cdots & R_{nn} \end{vmatrix}.$$

But for each $j = 0, \dots, n$ we have

$$\operatorname{ord}_j R_{hj}(z) \ge \sigma_j + \delta_{lj} - \delta_{0j}.$$

At this point we obtain some advantage from our general definition of the sequences $w$. For it is immediately evident that $\Psi(z|\sigma + \delta_l)$ divides the determinant in the sense that the hypotheses on the matrix $F$ entail that $\Psi(z|\sigma + \delta_l)$ divides $\det A(z|\rho, \sigma; l)$ in the ring $\mathbf{F}[z]$ of polynomials. Because both polynomials are monic polynomials of degree $|\sigma| = |\rho|$ it follows that

$$\det A(z|\rho, \sigma; l) = \Psi(z|\sigma + \delta_l).$$

Moreover, the coincidence of degrees makes it plain that for each $j = 0, \dots, n$ there is a nonempty subset $H(j) \subseteq \{0, \dots, n\}$ so that

$$\operatorname{ord}_j r_{hj}(z) = \sigma_j + \delta_{lj} - \delta_{0j}$$

for each $h \in H(j)$. Conversely, set $J(h) = \{j : h \in H(j)\}$. Our observations amount to $F$ having $J(h)$-normal approximation at $(\rho + \delta_h, \sigma + \delta_l)$. Whenever $J(h)$ is nonempty this is a nontrivial remark and since $\bigcup_h J(h) = \{0, \dots, n\}$ our observation is indeed nontrivial for some $h$. Then, as above, we may select an $l' \in J(h)$ and construct a nonsingular approximation matrix at $(\rho + \delta_h, \sigma + \delta_l; l')$; and so upwards.

The claim that every nonsingular approximation matrix lies on a normality zig-zag encompasses the fact just demonstrated, namely that, with $J$ and $J'$ both nonempty, above every $J$-normal point $(\rho, \sigma)$ there is a $J'$-normal point $(\rho + \delta_h, \sigma + \delta_l)$. Moreover, because the matrix $F$ is generic, some permutation matrix yields a nonsingular approximation matrix at $(0, 0)$. Finally, in the next section, we shall prove that if $F$ has $J$-normal approximation at $(\rho, \sigma)$ then, for some pair $(h, l)$, there is a nonsingular approximation matrix at $(\rho - \delta_h, \sigma - \delta_l)$. We shall also show that there are generalised normal points above every point.

## 4. Excess

Let $a$ be an approximation at $(\rho, \sigma)$. We define its *excess* to be the vector $\kappa = (\kappa_0, \dots, \kappa_n)$ of nonnegative integers $\kappa_j$ given by

$$\operatorname{ord}_j\left(\sum_i a_i f_{ij}\right) = \sigma_j - \delta_{0j} + \kappa_j .$$

Its $J$-excess $\kappa(J)$ is the $|J|$-tuple of $\kappa_j$, with $j \in J$. We set $\sum_{j\in J} \kappa_j = |\kappa(J)|$.

Suppose that there are essentially distinct $J$-best approximations at $(\rho, \sigma)$ each with $J$-excess $\kappa(J)$. Then, either all $\kappa_j$, $j \in J$, are infinite or a nontrivial linear combination of the approximations yields an approximation at $(\rho, \sigma)$ with $J$-excess strictly above $\kappa(J)$. In the generic case the series of any row of $F$ are linearly independent over the ring of polynomials, so the first possibility is excluded; the second possibility contradicts the maximality of a $J$-best approximation. Thus, up to normalisation, a $J$-best approximation at a point $(\rho, \sigma)$ is unique, regardless of its $J$-excess.

Let $a$ be a $J$-best approximation at $(\rho, \sigma)$ and with $J$-excess $\kappa(J)$. Fix some $l \in J$ and suppose that for $h = 0, \dots, n$ there are approximations to $F$ with order $\sigma + \delta_l$ at $\rho + \delta_h$ and with $J$-excess above $\kappa(J)$. Given existence, we may select normalised $J$-best such approximations $A_h = (A_{h0}, \dots, A_{hn})$, with the $A_{hh}$ monic. We denote the $(n+1)\times(n+1)$ matrix of the approximations by

$$A(z) = A(z|\rho, \sigma; J, l) = \left(A_{hi}(z|\rho, \sigma; J, l)\right).$$

We note that (once again since $|\kappa| \neq \infty$, because the functions in each row of $F$ are linearly independent over $\mathbf{F}[z]$) we have $\deg A_{hh} = \rho_h$ for each $h$. For if not, we would have an approximation to $F$ of order $\sigma$ at $\rho$ with $J$-excess strictly $J$-above $\kappa(J)$.

https://doi.org/10.1017/S1446788700033292 Published online by Cambridge University Press

We now study the determinants $\det A(z|\rho, \sigma; J, l)$ and find, as in the section above, that $\Psi(z|\sigma+\delta_l+\kappa(J))$ divides $\det A(z)$ in the ring of polynomials. But $\det A(z)$ is a monic polynomial of degree $|\rho|$, so this is a contradiction unless $|\kappa(J)|=0$. Hence, if $|\kappa(J)|\neq 0$, there cannot be, for every $h$, a $J$-best approximation to $F$ at $(\rho+\delta_h, \sigma+\delta_l)$ with $J$-excess $J$-above $\kappa(J)$. Thus, either the $J$-excess $\kappa(\rho, \sigma; J)$ at $(\rho, \sigma)$ is zero or, for some $h$, the $J$-excess $\kappa(\rho+\delta_h, \sigma+\delta_l; J)$ of a $J$-best approximation to $F$ at $(\rho+\delta_h, \sigma+\delta_l)$ is not $J$-above $\kappa(J)$. For such $h$, and provided that $\kappa_l\neq 0$, the approximation $a$ at $(\rho, \sigma)$ with excess $\kappa$, is already an approximation at $(\rho+\delta_h, \sigma+\delta_l)$ with excess $\kappa-\delta_l$.

Suppose now that $\kappa(\rho, \sigma; J)\neq 0$. Choose $l$ from amongst those $j\in J$ for which $\kappa_j > 0$. Then our argument, conducted with $\kappa(J)$ replaced by $\kappa(J)-\delta_l$, entails that, for some $h$, a $J$-best approximation to $F$ at $(\rho+\delta_h, \sigma+\delta_l)$ is already given by the $J$-best approximation $a$ at $(\rho, \sigma)$ and has $J$-excess $\kappa(J)-\delta_l$.

Repeating this argument $|\kappa(J)|$ times, we find a point $\rho'$ with $|\rho'| = |\rho|+|\kappa(J)|$ above $\rho$ and a corresponding order $\sigma'=\sigma+\kappa(J)$ above $\sigma$ at which the $J$-best approximation has zero $J$-excess. Thus $(\rho', \sigma')$ is a point above $(\rho, \sigma)$ at which $F$ has $J$-normal approximation. It follows that, as alleged, there are generalised normal points $\rho$ with $\rho_{\min}=\min_{0\leq i\leq n}\rho_i$ as large as we wish.

Conversely, suppose $F$ does not have generalised normal approximation above some point $\rho$. Recall that we described $F$ as generic if the series comprising each row of $F$ are linearly independent over the ring $\mathbf{F}[z]$ of polynomials with coefficients in $\mathbf{F}$ and if there is a permutation $\pi$ of $\{0, \dots, n\}$ so that $f_{0,\pi(0)}(w_{k_0,0})f_{1,\pi(1)}(w_{k_1,1})\cdots f_{n,\pi(n)}(w_{k_n,n})$ does not vanish for any selection of elements $w_{k(j),j}$ respectively from the sequences $w_j$. The second condition is the less interesting; if it does not hold for a given $F$, we can always restore as much of it as we need by replacing $F$ by a matrix with entries obtained by dividing the $f_{ij}$ by appropriate polynomials. Thus assume the second condition on $F$. Then if $F$ does not have generalised normal approximation above $\rho$, it follows from the preceding arguments that there is an approximation $a$ at $\rho$ with

$$\sum_i a_i(z)f_{ij}(z)=0 \quad \text{for some } j\in\{0, \dots, n\}.$$

Finally, suppose that $F$ has $J$-normal approximation at $(\rho, \sigma)\neq(0, 0)$ but there is no nonsingular approximation matrix at $(\rho-\delta_h, \sigma-\delta_l)$, for any pair $(h, l)$. This entails, for each pair $(h, l)$, that all entries of the excess vector $\kappa$ at $(\rho-\delta_h, \sigma-\delta_l)$ are positive. Thus approximations at $(\rho-\delta_h, \sigma-\delta_l)$ are approximations at $(\rho, \sigma)$. Hence, there is an approximation $a$ at $(\rho, \sigma)$ with $\deg a_i < \rho_i - 1$ for $i = 0, \dots, n$. But then for each pair $(h', l')$, the polynomials $(z - w_{h'l'})a_i$ yield an approximation at $(\rho + \delta'_h, \sigma + \delta'_l)$ with $\deg(z - w_{h'l'})a_{h'}(z) < \rho_{h'}$. This is impossible for all pairs $(h', l')$ because $F$ has $J$-normal approximation at $(\rho, \sigma)$. Of course, we assume all conventions necessary to maintain the validity of these claims when entries of $\rho$ or $\sigma$ are zero.

These remarks complete the proof of the normality zig-zag theorem.

## 5. Dual approximation

Suppose that $F$ is $J$-normal at $(\rho, \sigma)$ and that $A(z)$ denotes the approximation matrix at $(\rho, \sigma; l)$, for some $l \in J$. For each entry $A_{hi}$ of $A$ denote by $\mathfrak{A}_{hi}$ its cofactor. That is, for each pair $h, k$ we have the identities

$$\sum_{i=0}^{n} A_{hi}(z)\mathfrak{A}_{ki}(z) = \delta_{hk}\Psi(z|\sigma + \delta_l).$$

It is easy to compute that

$$\deg \mathfrak{A}_{ki} \leq |\sigma| - \rho_i - 1 \quad \text{if } k \neq i, \quad \text{and} \quad \deg \mathfrak{A}_{kk} = |\sigma| - \rho_k,$$

and to verify this against the identities. Furthermore, we see that the $\mathfrak{A}_{kk}$ are monic.

Using the terminology in which the author first learned the subject, we call the matrix

$$\mathfrak{A}(z|\rho, \sigma; l) = (\mathfrak{A}_{hi}(z|\rho, \sigma; l))$$

the *German approximation matrix* at $(\rho, \sigma; l)$; We now refer to the approximation matrix $A(z|\rho, \sigma; l)$ as the *Latin* approximation matrix.

Indeed, the German polynomials do solve an approximation problem: we say that a vector $\mathfrak{b} = (\mathfrak{b}_0, \dots, \mathfrak{b}_n)$ of polynomials $\mathfrak{b}_i$ is an $l$-German approximation to $F$ at $\rho$ of order $\sigma$ if for each $j$ and $i, s = 0, 1, \dots, n$,

$$\operatorname{ord}_j(\mathfrak{b}_i f_{sj} - \mathfrak{b}_s f_{ij}) \geq \sigma_j + \delta_{lj}$$

and

$$\deg \mathfrak{b}_i \leq |\sigma| - \rho_i.$$

Existence is a matter of $\sum(|\sigma| + 1 - \rho_i) = (n+1)(|\sigma| + 1) - |\rho|$ free coefficients which are to satisfy $n\sum(\sigma_j + \delta_{0j}) = n(|\sigma| + 1)$ independent linear conditions. So $F$ has an $l$-German approximation at $\rho$ of order $\sigma$ whenever $|\rho| \geq |\sigma| - 1$. Since we always take $|\sigma| = |\rho|$, the $l$-German approximation exists.

We will see that the rows of the German approximation matrix consist of the normalised $l$-German approximations at $(\rho - \delta_h, \sigma - \delta_0)$. To see this,

https://doi.org/10.1017/S1446788700033292 Published online by Cambridge University Press

let $\mathfrak{a}_k = (\mathfrak{a}_{k0}, \dots, \mathfrak{a}_{kn})$ be an $l$-German approximation at $(\rho - \delta_k, \sigma - \delta_0)$, set

$$\mathfrak{r}_{k,isj} = \mathfrak{a}_{ki}(z)f_{sj}(z) - \mathfrak{a}_{ks}(z)f_{ij}(z),$$

and consider sums

$$\begin{aligned} f_{sj}(z)\sum_{i=0}^{n} A_{hi}(z)\mathfrak{a}_{ki}(z) &= \sum_{i=0}^{n} A_{hi}(z)\big(\mathfrak{a}_{ki}(z)f_{sj}(z) - \mathfrak{a}_{ks}(z)f_{ij}(z)\big) \\ &\qquad + \mathfrak{a}_{ks}(z)\sum_{i=0}^{n} A_{hi}(z)f_{ij}(z) \\ &= \sum_{i=0}^{n} A_{hi}(z)\mathfrak{r}_{k,isj}(z) + \mathfrak{a}_{ks}(z)R_{hj}(z). \end{aligned}$$

But, for each $j$ and $i, s, k = 0, 1, \dots, n$,

$$\operatorname{ord}_j \mathfrak{r}_{k,isj}(z) \geq \sigma_j + \delta_{lj} - \delta_{0j}$$

and

$$\operatorname{ord}_j R_{hj}(z) \geq \sigma_j - \delta_{0j} + \delta_{lj}.$$

Given that the matrix $F$ is generic, there is a choice $s = s(j)$ for each $j = 0, 1, \dots, n$ so that we may collect these inequalities and conclude that the polynomial $\Psi(z|\sigma + \delta_l)$ divides the polynomial $\sum_{i=0}^{n} A_{hi}(z)\mathfrak{a}_{ki}(z)$ in the ring $\mathbf{F}[z]$ of polynomials.
However,

$$\deg \sum_{i=0}^{n} A_{hi}(z)\mathfrak{a}_{ki}(z) \leq \max_i(|\sigma| + \delta_{ki} - 1 + \delta_{hi} - 1).$$

It follows that

$$\sum_{i=0}^{n} A_{hi}(z)\mathfrak{a}_{ki}(z) = 0$$

unless both $h = k$ and $\mathfrak{a}_{kk}(z)$ attains its degree $|\sigma| - \rho_k$. In the latter case we may suppose that we have normalised so that each $\mathfrak{a}_{hh}(z)$ is monic. We obtain

$$\sum_{i=0}^{n} A_{hi}(z)\mathfrak{a}_{hi}(z) = \Psi(z|\sigma + \delta_l).$$

Notice that our decision to distinguish $\sigma_0$ in the definition of the Latin approximations forces the present definition of the dual approximations.

## 6. Transfer matrices

PROPOSITION. *Suppose $(\rho', \sigma')$ lies above $(\rho, \sigma; l)$ on a $J$-normality zig-zag. Then the quotient*

$$A(z|\rho', \sigma'; l')A(z|\rho, \sigma; l)^{-1}$$

*is a matrix with polynomial entries.*

PROOF. By duality we have

$$A(z|\rho', \sigma'; l')A(z|\rho, \sigma; l)^{-1} = A(z|\rho', \sigma'; l')\,{}^t\mathfrak{A}(z|\rho, \sigma; l)\big(\Psi(z|\sigma+\delta_l)\big)^{-1}.$$

The rest of the argument has much the same shape as that of the previous section, namely,

$$\begin{aligned} f_{sj}(z)\sum_{i=0}^{n} & A_{hi}(z|\rho', \sigma'; l)\mathfrak{A}_{ki}(z|\rho, \sigma; l) \\ &= \sum_{i=0}^{n} A_{hi}(z|\rho', \sigma'; l')\big(\mathfrak{A}_{ki}(z|\rho, \sigma; l)f_{sj}(z) - \mathfrak{A}_{ks}(z|\rho, \sigma; l)f_{ij}(z)\big) \\ &\qquad + \mathfrak{A}_{ks}(z|\rho, \sigma; l)\sum_{i=0}^{n} A_{hi}(z|\rho', \sigma'; l')f_{ij}(z) \\ &= \sum_{i=0}^{n} A_{hi}(z|\rho', \sigma'; l')\mathfrak{R}_{k,isj}(z|\rho, \sigma; l) \\ &\qquad + \mathfrak{A}_{ks}(z|\rho, \sigma; l)R_{hj}(z|\rho', \sigma'; l'); \end{aligned}$$

and for each choice of the subscripts

$$\operatorname{ord}_j \mathfrak{R}_{k,isj}(z|\rho, \sigma; l) \geq \sigma_j + \delta_{lj} - \delta_{0j}$$

and

$$\operatorname{ord}_j R_{hj}(z|\rho', \sigma'; l') \geq \sigma'_j - \delta_{0j} + \delta_{l'j}.$$

Given that the matrix $F$ is generic, there is a choice $s = s(j)$ for each $j = 0, 1, \ldots, n$ allowing us to conclude that the polynomial $\Psi(z|\sigma+\delta_l)$ indeed divides the polynomial $\sum_{i=0}^{n} A_{hi}(z|\rho', \sigma'; l')\mathfrak{A}_{ki}(z|\rho, \sigma; l)$ in the ring $\mathbf{F}[z]$ of polynomials.

Moreover,

$$\deg\sum_{i=0}^{n} A_{hi}(z|\rho', \sigma'; l')\mathfrak{A}_{ki}(z|\rho, \sigma; l) \leq \max_i(|\sigma| + \rho'_i - \rho_i + \delta_{ki} - 1 + \delta_{hi} - 1).$$

It follows that the $(h, k)$ entry of the matrix

$$A(z|\rho', \sigma'; l')A(z|\rho, \sigma; l)^{-1}$$

is a polynomial of degree at most

$$\max_i(\rho_i' - \rho_i + \delta_{ki} - 1 + \delta_{hi} - 1).$$

Special cases of these transfer matrices are explored *in extenso* by Mahler [7]. We note, without ado, that if $\max_i(\rho_i' - \rho_i)$ is at most 1 then the transfer matrix has constant entries other than on its diagonal where the entries may be constant or monic of degree 1. Thus it seems feasible to construct approximation matrices as products of transfer matrices of relatively simple shape. However, even in the special case $\sigma_1 = \cdots = \sigma_n = 0$ when only $\sigma_0$ is different from 0, the author is not aware of any constructions effected or estimations made in this manner.

## 7. Motivation for the present study

Consider the problem of constructing a polynomial $P(x, y)$ in two variables with prescribed vanishing at each of three points $(\alpha_x, \alpha_y)$, $(\alpha_x', \alpha_y')$ and $(\alpha_x'', \alpha_y'')$. The prescription will be of the following shape: $P$ is given to be of bi-degree $(m, n)$ and we require that

$$\frac{1}{i!j!}\frac{\partial^{i+j}}{\partial x^i \partial y^j}P(\alpha_x, \alpha_y) = 0 \quad \text{for } i = 0, 1, \dots, \tau_j - 1;\ j = 0, 1, \dots, n;$$

and similarly for the other two points. There is a solution provided that

$$3\sum \tau_j < (m+1)(n+1).$$

It is always appropriate to take

$$\tau_0 \geq \tau_1 \geq \cdots \geq \tau_n = 0.$$

To decrease data clutter it is convenient to apply linear fractional transformations to each variable transforming the data points to $(1, 1)$, $(\infty, \infty)$ and $(0, 0)$. The construction problem then acquires the shape

$$\frac{1}{j!}\frac{\partial^j}{\partial y^j}\sum_{i=0}^{n} A_i(x)(1-x)^{\tau_i}(1-y)^i \,\Big|_{y=0} = O(x^{\tau_j}) \quad \text{for } j = 0, 1, \dots, n,$$

with polynomials $A_i$ satisfying

$$\deg A_i(x) \leq m - \tau_i - \tau_{n-i} \quad \text{for } i = 0, 1, \dots, n.$$

One sees directly that

$$P(x, y) = \sum_{i=0}^{n} A_i(x)(1-x)^{\tau_i}(1-y)^i$$

has the prescribed vanishing at $(1, 1)$, whilst the conditions on the degree of the $A_i$ are the prescribed vanishing at $(\infty, \infty)$. The remaining problem is that of constructing the $A_i$ to provide the required vanishing at $(0, 0)$. That last problem is $\sum \tau_j$ linear conditions on $\sum m + 1 - \tau_i - \tau_{n-i}$ coefficients and is feasible provided that

$$\sum \tau_j < (m+1)(n+1) - 2\sum \tau_j ,$$

as at the beginning.

The required construction thus becomes a generalised simultaneous approximation problem of the shape

$$\sum_{i=0}^{n} \binom{i}{j} A_i(x)(1-x)^{\gamma_i} = O(x^{\tau_j}) \quad \text{for } j = 0, 1, \dots, n,$$

exactly as discussed in the body of this paper. It is of course a misfortune that, with $\gamma_i = \tau_i$ the given matrix, consisting of the functions

$$\binom{i}{j}(1-x)^{\gamma_i},$$

is quite degenerate. We have written $\gamma_i$ for $\tau_i$ to provide a more generic appearance and to suggest that the construction might well be attacked by viewing it as a limiting case of a generic approximation problem. In the case $n = 1$ this approach does yield the construction.

## References


[1] John Coates, 'On the algebraic approximation of functions I, II, III', *Nederl. Akad Wetensch. Proc. Ser. A* **69** (1966), 421–461; IV, *ibid* **70** (1967), 205–212.

[2] C. Hermite, 'Sur la fonction exponentielle', *Oeuvres* **3** (1873), 151–181.

[3] C. Hermite, 'Sur la généralisation des fractions continues algébriques', *Oeuvres* **4** (1893), 357–377.

[4] H. Jager, 'A multidimensional generalization of the Padé table', *Nederl. Akad. Wetensch. Proc. Ser. A* **67** (1964), 192–249.

[5] J. H. Loxton and A. J. van der Poorten, 'Multidimensional generalisations of the Padé table', *Rocky Mountain J. Math.* **9** (1979), 385–393.

[6] J. H. Loxton and A. J. van der Poorten, 'Arithmetic properties of automata: regular sequences', *J. für Math.* **330** (1988), 159–172.

[7] Kurt Mahler, 'Perfect systems', *Compositio Math.* **9** (1968), 95–166.


School of Mathematics, Physics, Computing and Electronics
Macquarie University
NSW 2109
Australia