Recursive algorithms for non normal padé tables

by A.BULTHEEL

to appear SIAM J. Appl. Math
vol 39, 1980

Report TW40, July 1978
Applied Mathematics and Programming Division
Katholieke Universiteit Leuven (Belgium)

1.ABSTRACT.

The Berlekamp-Massey algorithm [3,8] for minimal realization problems is a special case of the padé approximation problem. As a matter of fact, it computes among other polynomials the denominators of the elements of the padé table that are on the descending diagonal $\{[0/1],[1/2],\ldots,[k/k+1],\ldots\}$ as far as they exist and this algorithm works for nonnormal padé tables too. This algorithm does not seem to be well known in padé approximation literature. It is not very difficult to generalize this algorithm so as to compute the other padé approximants of a nonnormal table. Some variants will lead to a generalization of the algorithm of Brezinski [1], which computes the descending diagonals of a normal padé table, and of the algorithms of Watson [3,4,10] and Thacher [3,4] that compute the descending staircases of normal tables. This work is the dual of Cordellier's [6], deriving similar results for ascending diagonals and staircases.

We suppose that the reader is somewhat familiar with padé approximation and with the Berlekamp-Massey algorithm.

2. INTRODUCTION

We recall some basic definitions and some well known properties without proof for further reference.

Corresponding to a formal power series $\sum_{k=0}^{\infty} c_k z^k$ we define for all $L,M \geqslant 0$ a *padé form*

$$(L/M)_N = \frac{P^{(L/M)}(z)}{Q^{(L/M)}(z)}$$

as any couple of polynomials $P^{(L/M)}$, $Q^{(L/M)}$ (if it exists) satisfying

$$(PF)_N \begin{cases} 1.\ d^\circ\ P^{(L/M)} = L \\ 2.\ d^\circ\ Q^{(L/M)} = M \quad \text{(but the degree need not be strict)} \\ 3. P^{(L/M)} \text{ and } Q^{(L/M)} \text{ are relative prime.} \\ 4.\ Z^{(L/M)} = Q^{(L/M)} f - P^{(L/M)} = r_N z^N + r_{N+1} z^{N+1} + \ldots \\ 5.\ + \text{some normalization.} \end{cases}$$

If this couple does not exist then we set by convention $(L/M)_N = \text{nill}$.

If $(L/M)_N \neq \text{nill}$ and $r_N \neq 0$ then N is called its *order*.

$Z^{(L/M)}$ is called the *residual*.

Normalization can be chosen in several ways. E.g. making P or Q monic or comonic. Here we will choose as normalizing condition $Q^{(L/M)}(0) = 1$.

Strictly spoken, there is a difference between e.g.

$$\frac{1}{1} \quad \text{and} \quad \frac{1+0z}{1+0z+0z^2}$$

because the first is an element of the form $(0/0)_N$ and the second is of the form $(1/2)_N$. We will go rather loosely over this subtle difference and regard them as equal.

If the order of a padé form $(L/M)_N$ is at least L+M+1 then it is called a *padé approximant* and it is denoted by $[L/M]$. As is well known: a padé approximant is uniquely defined.

$\{L/M\}$ is the set of all padé forms of some arbitrary order.

$$\{L/M\} = \{(L/M)_N \mid N = 0,1,2,\ldots\}$$

Its elements are nill and all the rational forms that are reduced, normalized and with d° numerator $\leqslant L$ and d° denominator $\leqslant M$

Property 1

$$\{I/J\} \subset \{L/M\} \quad \Leftrightarrow \quad I \leqslant L \text{ and } J \leqslant M \ .$$

□

A *padé table* is a tableau containing in its L-th row and M-th column the element [L/M].
A padé table is *normal* if all [L/M] are different from nill (and are different from each other).
The structure of a nonnormal padé table is well known.(see e.g. the example in appendix 1). It is devided into square *blocks* of size n x n ($n \geq 1$) . If $n > 1$ then the block is a *singular block* and a singular block contains up and above its antidiagonal non-nill padé approximants that are all equal to the left top element. Below the antidiagonal all the entries are nill.
The left top element of a (singular) block will be called the *indicator* of the (singular) block and of its elements.
The padé table of a series

$$c_n z^n + c_{n+1} z^{n+1} + \ldots \qquad (n > 0 \ , \ c_n \neq 0)$$

contains such a n x n block in its left top corner. To the right of it we find a band of size $n \times \infty$ with nill all over. To avoid difficulties with this band, we suppose that the leading coefficient $c_0 \neq 0$ in the following. This is no restriction really since for $L \geq n$

$$[L/M]_{z^n f(z)} = z^n [L-n/M]_{f(z)}$$

Property 2

$c_0 \neq 0 \Rightarrow$ [k/0] and [0/k] are non nill for all k □

The *order* N of {L/M} is the highest possible order that can be obtained by an element of {L/M}.
A padé form $(L/M)_N \in \{L/M\}$ having this maximal order is called an *order representant* of {L/M}. This definition makes sense only if not all elements of {L/M} are nill. By property 1 we have that {L/M} will contain among others the entries of the padé table that are on or within the rectangle that has as a diagonal [0/0] - [L/M], especially [L/M] [L/0] and [0/M]. The last two being non nill by property 2.

Property 3

The order of {L/M} is at least max(L,M) + 1 □

In appendix 2 we will prove:

Property 4

The order of {L/M} is the order of the indicator of [L/M].

□

If [L/M] = nill then we call {L/M} *degenerate* (i.e. (L/M) indicates the position in a singular block below the antidiagonal.)

Property 5

{L/M} is non degenerate iff its order is at least L+M+1

□

Since a padé approximant is uniquely defined, it follows from property 4 that:

Property 6

The order representant of {L/M}is unique and is some padé approximant

□

It is known that in the normal case the elements of a descending staircase T_k can be computed recursively by the following relations:
Let $R^{(1)}$, $R^{(2)}$ and $R^{(3)}$ be three successive elements from a staircase and let

$$S^{(i)} = \alpha P^{(i)} + \beta Q^{(i)}$$

α,β arbitrary expressions

(e.g. $(\alpha,\beta) = (1,0)$ iff $S^{(i)} = P^{(i)}$; $(\alpha,\beta) = (0,1)$ iff $S^{(i)} = Q^{(i)}$ and $(\alpha,\beta) = (-1,f(z))$ iff $S^{(i)} = Z^{(i)}$.)

Then we have [3,4]

$$S^{(3)} = S^{(2)} - r_2 r_1^{-1} z S^{(1)} \qquad (1)$$

where r_i are the leading non zero coefficients of $Z^{(i)}$.
These coefficients can be computed by

a) previous recursions of the form (1) for $S^{(i)} = Z^{(i)}$
b) by expressions of the form (2)

$$r_2 = \sum_{\omega} b_{\omega}^{(2)} c_{N-\omega}$$

where $Q^{(2)} = \sum_{\omega} b_{\omega}^{(2)} z^{\omega}$ and $Z^{(2)} = O(z^N)$

The algorithm of Watson [3,4,10] uses relations (1) and (2b) and the Thacher algorithm [3,4] uses (1) and (2a) in an unnormalized form.
There also exist similar recursions between adjacent elements of a descending diagonal

$$S^{(3)} = (1 + F_1 z) S^{(2)} + F_2 z^2 S^{(1)}$$

with

$$F_1 = r_2' r_2^{-1} - r_1' r_1^{-1} \quad ; \quad F_2 = - r_1 r_2^{-1}$$

where $Z^{(i)} = r_i z^{N_i} + r_i' z^{N_i+1} + \ldots$ is of order N_i .

These r_i and r_i' can again be expressed in terms of $b_\omega^{(i)}$ and c_ω.
This is known as Brezinski's algorithm [1].

To generalize these algorithms to the nonnormal case, we can see that the elements of such a nonnormal staircase (diagonal) may be nill or repeat themselves. Moreover the order of the approximant and the degrees of numerator and denominator are not precisely known a priori. So we will need a more complicated notion of a staircase, order and degree. This is done in the next section.

3.Diagonals,Staircases, Order and Degree

A *diagonal* $\mathcal{D}_k$ is an ordered set consisting of the elements

$$\{ \{k/0\},\{k+1/1\}, \ldots ,\{k+\ell/\ell\}, \ldots \} \text{ for } k \geqslant 0$$

and

$$\{ \{0/k\},\{1/k+1\}, \ldots ,\{\ell/-k+\ell\}, \ldots\} \text{ for } k < 0$$

Selecting only one padé form (the padé approximant) for every $\{L/M\} \in \mathcal{D}_k$, we will get an ordered set of padé forms (padé approximants) which will be refered to as $\hat{\mathcal{D}}_k (\hat{D}_k)$.
If we get rid of all the nill elements and the duplicate elements in $\hat{\mathcal{D}}_k$ ($\hat{D}_k$) we will denote this as $\mathcal{D}_k$ (D_k) .
This means that D_k contains exactly one element for every (singular) block it passes trough viz. its indicator.

By a (descending) staircase $\mathcal{T}_k$ we mean the ordered set

$$\{ \{k/0\},\{k+1/0\},\{k+1/1\},\{k+2/1\},\ldots,\{k+\ell/\ell\},\{k+\ell+1/\ell\},\ldots\} \text{ if } k \geqslant 0$$

and

$$\{ \{0/-k\},\{0/-k+1\},\{1/-k+1\}, \ldots,\{\ell/-k+\ell\},\{\ell/-k+\ell+1\},\ldots \} \text{ if } k < 0$$

There is clearly a discontinuity in this definition for k = 0 because the staircase

$$\{ \{0/0\},\{0/1\},\{1/1\},\{1/2\}, \ldots\}$$

is not covered. So let us call this one $\mathcal{T}_{-0}$ and the previously defined $\mathcal{T}_0$ will be called $\mathcal{T}_{+0}$ from now on. Thus we will always suppose that the indices of $\mathcal{T}_k$ will be taken from the ordered set $\{\ldots,-2,-1,-0,+0,1,2,\ldots\}$ and $\mathcal{T}_{-0} \neq \mathcal{T}_{+0}$ but +0 and -0 will coincide in every other appearance.
We will often give a proof of some property for $k \geqslant +0$ only because mutatis mutandi the proof for $k \leqslant -0$ will be the same.
Naturaly $\hat{\mathcal{T}}_k, \mathcal{T}_k$, $\hat{T}_k$ and T_k are defined similarly to $\hat{\mathcal{D}}_k, \mathcal{D}_k, \hat{D}_k$ and D_k.
As is seen from the definition $\mathcal{T}_k$ is an interleaving of $\mathcal{D}_k$ and $\mathcal{D}_{k+1}$ for $k \geqslant +0$ and of $\mathcal{D}_k$ and $\mathcal{D}_{k-1}$ for $k \leqslant -0$.
Because we do not want to make this double reference always, we will refer to $\mathcal{D}_k$ as the *inner* diagonal of $\mathcal{T}_k$ and to the other as the *outer* one.
An element $\{L/M\} \in \mathcal{T}_k$ will always be an element of a neighbouring staircase too, which is closer to $\mathcal{D}_0$ or not, depending on whether

{L/M} is on the inner or on the outer diagonal of T_k.

If {L/M} $\in T_k$ then we will say that its *degree* is

$$M + k + 1 \qquad \text{if} \quad k \geq +0$$
$$L - k + 1 \qquad \text{if} \quad k \leq -0$$

We see that the elements of T_k have degrees

$k+1, k+1, k+2, k+2, \ldots, k+\ell+1, k+\ell+1, \ldots$ for $k \geq +0$

$-k+1, -k+1, -k+2, -k+2, \ldots, -k+\ell+1, -k+\ell+1, \ldots$ for $k \leq -0$

The following is trivial:

Property 7

$d \geq |k| + 1$ iff T_k contains exactly two elements of degree d. □

That $R \in T_k$ has degree d means that its elements $R = P/Q \in R$ that are non nill can be represented by

$$P = \sum_{j=0}^{d} a_j z^j \quad ; \quad Q = \sum_{j=0}^{d-k-1} b_j z^j \qquad \text{for} \quad k \geq +0$$

and

$$P = \sum_{j=0}^{d-|k|-1} a_j z^j \quad ; \quad Q = \sum_{j=0}^{d} b_j z^j \qquad \text{for} \quad k \leq -0$$

If the degree is strict for P and/or for Q then we say that the padé form R has *degree* d.
If $R \in R \in T_k$ has the same degree as R, then we call it a *degree representant* of R.

Remark that the degree of {L/M} is dependent on the staircase it is considered to be an element of. Since {L/M} is always on two neighbouring staircases, it has also two possible degrees, but it will be clear from the context which one is meant.

By writing a subscript for R: $R_d \in T_k$ we mean that R_d has degree d. There are two possible elements that can be indicated by R_d : one that is on the inner diagonal and one that is on the outer diagonal. If we want to make a difference between them then we write R_d^I and R_d^O respectively.

By property 7 R_0 will never be a member of T_k but for notational comfort we will put

$$R_{+0} = \{\tfrac{1}{0}\} \in T_k \quad \text{for} \quad k \geq +0 \quad \text{and} \quad R_{-0} = \{\tfrac{0}{1}\} \in T_k \quad \text{for} \quad k \leq -0$$

Because $c_0 \neq 0$ its order is always 0.

Next we will investigate some relations between the order and the degree of some $R \in T_k$.

It follows from the definition that the set R grows with its degree so we have clearly

Property 8

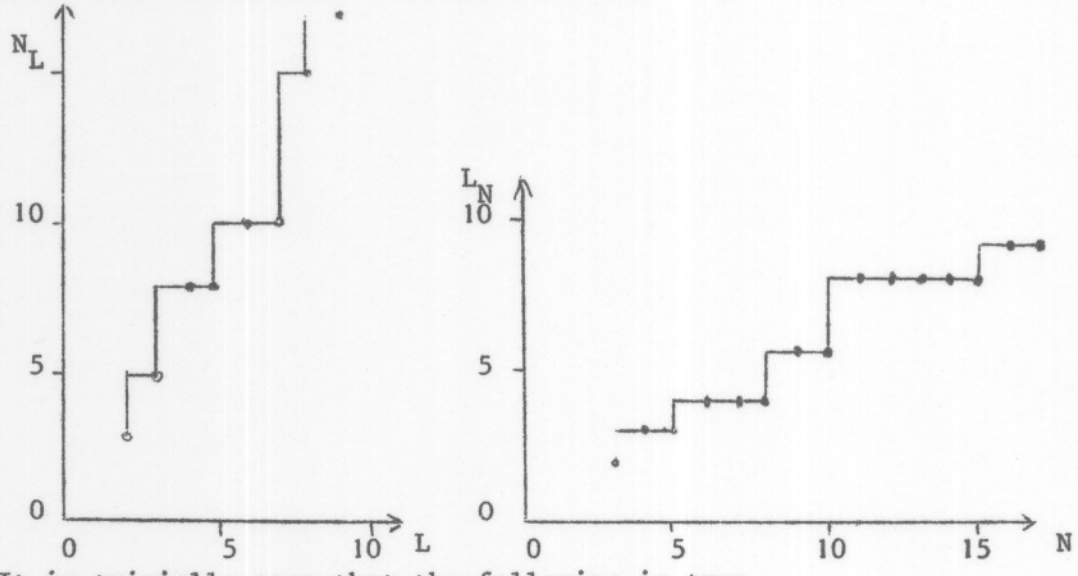
The order of $R_d \in T_k$ is nondecreasing with its degree. □

We will now need the following concepts:

N_L = the highest possible order obtainable by an element of degree $\leq L$ (this is the order of R^0_L) and its inverse is

L_N = the lowest possible degree allowing an order $\geq N$ approximation.

These are clearly not 1-1 relations: e.g.



It is trivially seen that the following is true.

Property 9

$$L_{N+1} \geq L_N \geq L_0 = |k|$$ □

Property 10

If $R \in R_{L_N} \in T_k$ is of order $\geq N$ then it is a degree representant of R_{L_N} (i.e. it is of degree L_N).

Proof:

Clearly if the degree of R were $d < L_N$ then L_N would not be the lowest degree allowing an order $\geq N$. □

Corollary 1

The order representant of $R^0_{L_N} \in T_k$ is a degree representant of $R^0_{L_N}$. □

Property 11

If $L_N < L_{N+1}$ then $R_{L_N} \in T_k$ has an order which is not higher than N.

The order of $R^O_{L_N}$ is exactly N iff $L_N < L_{N+1}$.

Proof:

If the order of R_{L_N} were $\geq N+1$ then we would have that $L_{N+1} \leq L_N$. This proves the first part. Since $R^I_{L_N} \subset R^O_{L_N}$ we must have by definition of L_N that the order of $R^O_{L_N}$ is at least N. This and the first part give the necessary condition.

By property 1 we know that $R^O_{L_N}$ contains all rational forms of a degree $\leq L_N$, so the order is the highest obtainable by a degree $\leq L_N$. This means that $L_{N+1} > L_N$. This proves the sufficient condition. □

Corollary 2

If $L_N < L_{N+1}$ and $R \in R^O_{L_N} \in T_k$ has an order N then R is the order representant of $R^O_{L_N}$. □

Property 12

The order of $R^O_{L_N}$ is at least $2\,L_N - |k| - 1$.

It is at least $2\,L_N - |k|$ if $R^O_{L_N}$ is non degenerate.

Proof : (for $k \geq +0$)

$R^I_{L_N}$ and $R^O_{L_N}$ can not be both degenerate because then its order representant would be of a lower degree and have an order at least N. This contradicts the definition of L_N.

Now $R^I_{L_N} = \{L_N - 1 \;/\; L_N - k - 1\}$ and $R^O_{L_N} = \{L_N / L_N - k - 1\}$

If $R^I_{L_N}$ is non degenerate then its order is at least $2\,L_N - k - 1$ (property 5) and the order of $R^O_{L_N}$ is not less than the order of $R^I_{L_N}$ (property 8 and 9) . If $R^O_{L_N}$ is non degenerate then its order is at least $2\,L_N - k$ (property 5). □

Property 13

If R exists having degree L_N and an order $N' \geq N$ then $L_{N'} = L_N$.

Proof :

The existence of R clearly implies that $L_{N'} \leq L_N$. It can not be less because of property 9. □

Proposition 1 (Massey)

If there exists a padé form $R \in \mathcal{R} \in T_k$ of degree p, having an order $N \geq p+1$ and if there exists a padé form S having degree q and an order $M \geq N+1$ then $q \geq N + |k| + 1 - p$

Proof : (for $k \geq +0$)

Since R is of degree p and of order N we have by the 4-th condition of $(PF)_N$ that

$$\left(\sum_{j=0}^{\infty} c_j z^j\right)\left(\sum_{j=0}^{p-k-1} b_j^R z^j\right) - \left(\sum_{j=0}^{p} a_j^R z^j\right) = d_N z^N + \ldots \qquad d_N \neq 0$$

or equivalently

$$\begin{aligned} \sum_{j=0}^{p-k-1} b_j^R c_{\ell-j} &= a_\ell^R && \ell = 0,1,\ldots,p \\ &= 0 && \ell = p+1,\ldots,N-1 \\ &\neq 0 && \ell = N \end{aligned}$$

By the normalisation is $b_0 = 1$ and thus

$$\begin{aligned} \sum_{j=1}^{p-k-1} b_j^R c_{\ell-j} &= -c_\ell && \ell = p+1,\ldots,N-1 \\ &\neq -c_\ell && \ell = N \end{aligned} \tag{4}$$

Similarly:

$$\sum_{j=1}^{q-k-1} b_j^S c_{\ell-j} = -c_\ell \qquad \ell = q+1,\ldots,N \tag{5}$$

Now suppose $q < N+k+1-p$ then taking (4) for $\ell = N$ and filling in with (5) gives:

$$\begin{aligned} -c_N \neq \sum_{j=1}^{p-k-1} b_j^R c_{N-j} &= \sum_{j=1}^{p-k-1} b_j^R \left(-\sum_{i=1}^{q-k-1} b_i^S c_{N-j-i}\right) \\ &= \sum_{j=1}^{q-k-1} b_i^S \left(-\sum_{j=1}^{p-k-1} b_j^R c_{N-i-j}\right) \\ &= \sum_{i=1}^{q-k-1} b_i^S c_{N-i} = -c_N \end{aligned} \tag{6}$$

Since by supposition $N-q+k-1 \geq p+1$ the last two steps are verified by (4). (6) is a contradiction and this proves the proposition.

□

Corollary 3

1. If $L_{N+1} > L_N$ then $L_{N+1} \geq N+|k|+1-L_N$ (7)

2. If R of degree L_N and of order $N \geq L_N + 1$ exists then

$$L_{N+1} \geq \max(L_N, N+|k|+1-L_N) \quad (8)$$

Proof :

1.By corollary 1 we know that the order representant of $R^0_{L_N}$ has degree L_N and by property 11 we know that its order is exactly N. The order representant of $R^0_{L_{N+1}}$ is of an order at least N+1 and of degree L_{N+1}.

Taking these order representants as R and S in proposition 1 we find the desired inequality.

2.Since $R^0_{L_{N+1}}$ has an order representant of degree L_{N+1} (corollary 1) and of an order at least N+1. The proposition is again applicable. Together with property 9 this gives (8).

□

4.Algorithms

Our aim will be to construct a staircase T_k with elements of strictly growing order, keeping the degree as low as possible, thus not increasing the degree L until the element of highest possible order $N' = N_L$ is found. Then we increase the degree from L to $L_{N'+1}$ etc...
We directly formulate the algorithm and prove its correctness.

Theorem 1

For some integer $k \in \{ \ldots,-2,-1,-0,+0,1,2,\ldots\}$ construct the staircase T_k consisting of the elements $R^{(i)}$ $i = 1,2,\ldots$ by the following algorithm A

A:

Initialisation:

$d_0 = 0$; $R^{(0)} = \frac{1}{0}$ $(k \geqslant +0)$ or $\frac{0}{1}$ $(k \leqslant -0)$; $N'_0 = 0$;

$r_0 = -1$ $(k \geqslant +0)$ or c_0 $(k \leqslant -0)$;

$d_1 = |k|$; $R^{(1)} = [\,k/0\,]$ $(k \geqslant +0)$ or $[\,0/-k\,]$ $(k \leqslant -0)$;

N'_1 = the order of $R^{(1)}$;

r_1 = the leading non zero coefficient of $Z^{(1)}$;

Recursively for $i = 2,3,\ldots$ do the following:

1. Select ρ_i by the following criterium:
   $\mathcal{K}$: ρ_i is some ρ such that $N'_\rho + |k| + 1 - d_\rho = d_{\rho+1} = d_{\rho+2} = \ldots = d_{i-1}$;
2. Form
$$P^{(i)} = P^{(i-1)} - F_i \, P^{(\rho_i)} z^{\mu_i}$$
$$Q^{(i)} = Q^{(i-1)} - F_i \, Q^{(\rho_i)} z^{\mu_i}$$
   with $F_i = r_{i-1} \, r_{\rho_i}^{-1}$ and $\mu_i = N'_{i-1} - N'_{\rho_i}$
3. For $R^{(i)} = \frac{P^{(i)}}{Q^{(i)}}$ compute
   N'_i = the order of $R^{(i)}$
   r_i = the leading non zero coefficient of $Z^{(i)}$
4. Set
$$N_i = N'_{i-1} + 1 \quad \text{and} \quad d_i = \max\,(d_{i-1}\,,\, N_i + |k| - d_{i-1})$$

end of algorithm A

Then we have:

1. $N_i' \geqslant N_i \qquad i = 1,2,\ldots$
2. $R^{(i)}$ is of degree $d_i < N_i' \qquad i = 1,2,\ldots$
3. The outer diagonal of T_k is a part of $\mathcal{T}_k$. It is formed by the elements $R^{(i)}$ for which $d_i < d_{i+1}$.

Comment :

If the theorem is proved then two possible variants of criterium $\mathcal{K}$ are:

$$\mathcal{K}_1 : \rho_i = \max_{\rho \leqslant i-1} \{\rho \mid d_\rho < d_{\rho+1}\}$$

or

$$\mathcal{K}_2 : \rho_i = \max_{\rho < i-1} \{\rho \mid d_\rho \leqslant d_{\rho+1}\}$$

Proof :

The theorem is true for i = 1,2. For the inductionstep we have:

1. Because

$$Z^{(i-1)} = r_{i-1} z^{N_{i-1}'} + \ldots$$

$$Z^{(\rho_i)} = r_{\rho_i} z^{N_{\rho_i}'} + \ldots$$

and

$$Z^{(i)} = Z^{(i-1)} - r_{i-1} r_{\rho_i}^{-1} z^{N_{i-1}' - N_{\rho_i}'} Z^{(\rho_i)}$$

$R^{(i)}$ must be at least of order $N_{i-1}' + 1 = N_i$.

2. Let us determine the degree of $R^{(i)}$.

By construction we have:

degree $R^{(i)} \leqslant \max(\text{degree } R^{(i-1)}, \mu_i + \text{degree } R^{(\rho_i)})$

and by induction hypothesis:

$$\text{degree } R^{(i-1)} = d_{i-1}$$

$$\text{degree } R^{(\rho_i)} = d_{\rho_i}$$

and criterium $\mathcal{K}$ gives:

$$d_{i-1} = d_{i-2} = \ldots = d_{\rho_i+1} = N_{\rho_i}' + |k| + 1 - d_{\rho_i}$$

or

$$d_{\rho_i} = N_{\rho_i}' + |k| + 1 - d_{\rho_{i+1}}$$

so that

$$d_{\rho_i} + \mu_i = N'_{i-1} + |k| + 1 - d_{i-1}$$

So we have that:

$$\text{degree } R^{(i)} \leqslant \max(d_{i-1}\ ,\ N'_{i-1} + |k| + 1 - d_{i-1}) = d_i$$

Since its order is at least N_i we have

$$L_{N_i} \leqslant \text{degree } R^{(i)} \leqslant d_i \tag{9}$$

By induction hypothesis again we have that degree $R^{(i-1)} = d_{i-1} = L_{N_{i-1}}$ and by property 13 this is also $L_{N'_{i-1}}$. Its order is $N'_{i-1} > L_{N'_{i-1}}$ so because of corollary 3.2) we have that

$$L_{N'_{i-1}+1} = L_{N_i} \geqslant$$

$$\geqslant \max(L_{N'_{i-1}}\ ,\ N'_{i-1} + |k| + 1 - L_{N'_{i-1}}) = \max(L_{N_{i-1}}, N'_{i-1} + |k| + 1 - L_{N_{i-1}}) \tag{10}$$

(9) and (10) then imply that $L_{N_i} = \text{degree } R^{(i)} = d_i$.
Point 2 is proved if $d_i < N'_i$.
By (10) we know that we have one of the two :

a) $d_i = L_{N_i} = L_{N'_{i-1}} < N'_{i-1}$ (by induction)

b) $d_i = N'_{i-1} + |k| + 1 - L_{N_{i-1}} \leqslant N'_{i-1}$ (property 7).

Thus in both cases we have $d_i \leqslant N'_{i-1} < N'_i$ while N'_j is strictly increasing.

3. Take those $R^{(\ell)}$ such that $d_\ell < d_{\ell+1}$. By property 11 we then have that the order or R_{d_ℓ} is at most N_ℓ but its order is at least N_ℓ as proved in the first part of this theorem. We thus have that $N_{\ell+1} = N'_\ell + 1 = N_\ell + 1$. Corollary 2 then says that $R^{(\ell)}$ is the order representant of $R^0_{L_{N_\ell}}$. Since N_ℓ is strictly increasing we have that $R^{(\ell)}$ can not be duplicated and it must be the element of T_k having the order N_ℓ. Conversely all elements of T_k must be part of T_k because we can not jump over one of them by corollary 3.

□

Remark

We could write down directly the recursion between elements $R^{(\ell)}$ that are on the outer diagonal of T_k.

Set as before $S = \alpha P + \beta Q$ and let

$$d_m < d_{m+1} = d_{m+2} = \ldots = d_i < d_{i+1} = d_{i+2} = \ldots = d_j < d_{j+1}$$

for simplicity let us take criterium $\mathcal{H}_2$ then

$$\rho_\ell = m \qquad \text{for} \qquad \ell = m+2, \ldots, i+1$$

and

$$\rho_\ell = i \qquad \text{for} \qquad \ell = i+2, \ldots, j+1$$

$$\begin{aligned} S^{(j)} &= S^{(j-1)} - F_j z^{\mu_j} S^{(i)} \\ &= S^{(j-2)} - (F_{j-1} z^{\mu_{j-1}} + F_j z^{\mu_j}) S^{(i)} \\ &= \ldots \\ &= S^{(i+1)} - (F_{i+2} z^{\mu_{i+2}} + \ldots + F_j z^{\mu_j}) S^{(i)} \\ &= S^{(i)} - F_{i+1} z^{\mu_{i+1}} S^{(m)} - (F_{i+2} z^{\mu_{i+2}} + \ldots + F_j z^{\mu_j}) S^{(i)} \\ &= (1 - F_{i+2} z^{\mu_{i+2}} - \ldots - F_j z^{\mu_j}) S^{(i)} - F_{i+1} z^{\mu_{i+1}} S^{(m)} \end{aligned} \qquad (11)$$

(11) can be thought of as the forward recursion for the evaluation of a continued fraction whose convergents are the successive outer diagonal elements of T_k. E.g. for $k \geqslant +0$:

$$[k/0] + \sum_\ell \left| \frac{\phi_\ell(z)}{\psi_\ell(z)} \right| - \left| \frac{F_{i+1} z^{\mu_{i+1}}}{1 - F_{i+2} z^{\mu_{i+2}} - \ldots - F_j z^{\mu_j}} \right| + \ldots$$

Such a relation is very much in the style of the continued fractions considered in [5] and under certain conditions are exactly those mentioned in [2].

Proposition 2

Applying algorithm A in the normal case we have for $i \geqslant 2$

1) $d_{i-1} \leqslant d_i \leqslant d_{i-1} + 1$ (i.e. $N'_{i-1} + |k| \leqslant 2d_{i-1}$)

2) if $d_i = d_{i-1}$ then $N'_i = N_i = 2d_i - |k|$

$$\mu_i = 1$$

$$d_{i+1} = d_i + 1$$

if $d_i = d_{i-1} + 1$ then $2d_i - |k| - 1 \leqslant N'_i \leqslant 2d_i - |k|$

Comment :

This means that you can have at most two consecutive degrees that are equal and (as in the general case) the last element of a certain degree will have the maximum possible order that is obtainable by that degree.

Proof : (for $k \geq +0$)

The proof is by induction and the verification of the initial step is left to the reader.

1.Because $d_\ell = \max(d_{\ell-1}, N'_{\ell-1} + k + 1 - d_{\ell-1})$ we will find by filling in the possible values for $N'_{\ell-1}$ that are allowed by the induction hypothesis 1) will be proved.

2.a) Now the conjecture about the orders:

R_d must be one of {d-1/d-k-1} or {d/d-k-1} and these have orders that are exactly 2d-k-1 and 2d-k respectively.

We have by the first part of the proposition i) or ii)

i) $d_{\ell-1} = d_{\ell-2}$ then by induction $d_\ell = d_{\ell-1}+1$ and thus

$$2d_\ell - k = 2d_{\ell-1} - k + 2 = N_{\ell-1} + 2 = N_\ell + 1$$

and because of 1)

$$N'_\ell \leq 2d - k = N_\ell + 1$$

and of course

$$N_\ell \leq N'_\ell$$

ii) $d_{\ell-1} = d_{\ell-2} + 1$ then because $d_\ell = \max(d_{\ell-1}, N'_{\ell-1}+k+1-d_{\ell-1})$ we have that if $N'_{\ell-1} = 2d_{\ell-1}-k$ then $d_\ell = d_{\ell-1}+1$

and if $N'_{\ell-1} = 2d_{\ell-1}-k-1$ then $d_\ell = d_{\ell-1}$

Filling in the possible values for $N'_{\ell-1}$ and $d_{\ell-1}$ proves the conjecture as in case i)

2.b) The conjecture about μ_i follows by simple inspection.

□

Suppose that in the normal case we have that

$$d_{i-2} < d_{i-1} = d_i < d_{i+1} = d_{i+2}$$

then if algorithm A is used with criterium $\mathcal{H}_1$

$$\begin{aligned} S^{(i+2)} &= S^{(i+1)} - r_{i+1}\, r_i^{-1}\, z\, S^{(i)} \\ &= S^{(i)} - r_i\, r_{i-2}^{-1}\, z^2\, S^{(i-2)} - r_{i+1}\, r_i^{-1}\, z\, S^{(i)} \\ &= (1 - r_i r_{i-1}^{-1} z)\, S^{(i)} - r_i r_{i-1}^{-1}\, z^2\, S^{(i-2)} \end{aligned} \qquad (12)$$

Expressing r_{i+1} as a function of residual coefficients of $Z^{(i-2)}$ and $Z^{(i)}$ we find that this is the method of Brezinski. Expression (11) can be seen as a generalization of (12) to the nonnormal case. If algorithm A is used with criterium $\mathcal{K}_2$ then we find the Watson-Thacher recursions.

We will now specify $\mathcal{K}_3$ as a special form of criterium $\mathcal{K}$ and prove that then not only the outer diagonal of T_k is constructed, but also the inner one.

$\mathcal{K}_3$: for i = 2,3,...

use $\mathcal{K}_1$: $\rho_i = \max_{\rho < i-1} \{\rho | d_\rho < d_{\rho+1}\}$ to select ρ_i

until $d_i < d_{i+1}$. From then on

use $\mathcal{K}_2'$: $\rho_i = \max_{\rho < i-1} \{\rho | d_\rho \leq d_{\rho+1}$ and $R^{(\rho)} \in T_k\}$

until $N'_{j-1} \geq 2d_{j-1} - |k| - 1$. From then on use again $\mathcal{K}_1$ etc...

Theorem 2

If in algorithm A we take $\mathcal{K}_3$ as a special case of $\mathcal{K}$ and call the corresponding algorithm A3 then:

1. The statements of theorem 1 are valid for A3.
2. The last elements $R^{(i)}$ for which ρ_i is selected by $\mathcal{K}_1$ before a switch to $\mathcal{K}_2'$ form the outer diagonal of T_k.
3. The last elements $R^{(i)}$ for which ρ_i is selected by $\mathcal{K}_2'$ before a switch to $\mathcal{K}_1$ form the inner diagonal of T_k.

Proof : (for $k \geq +0$)

Clearly we only have to prove 3). Therefore we will first prove the following lemma:

Lemma

If ρ_i is selected by $\mathcal{K}_2'$ then $P^{(i)}$ (the numerator of $R^{(i)}$) is of degree $d^\circ P^{(i)} \leq d_i - 1$ (i.e. $R^{(i)} \in R^I_{d_i}$).

Proof : (for $k \geq +0$)

$R^{(1)} = [k/0]$ is of degree k+1 and $d^\circ P^{(1)} \leq k$. Suppose that

$$d_2 = d_3 = \ldots = d_i < d_{i+1}$$

then $\rho_2, \ldots, \rho_i$ are selected by $\mathcal{K}_1$ and ρ_{i+1} by $\mathcal{K}_2'$.

$\frac{1}{0}$ is the first member of T_k (outer diagonal) and [k/0] is the second (inner diagonal) and by theorem 1 $R^{(i)}$ must be the third (outer diagonal). $\mathcal{H}_2'$ then gives $\rho_{i+1} = 1$. By construction we then have that $d°P^{(i+1)} \leqslant d_{i+1}-1$. If $N'_{i+1} < 2d_{i+1} - k - 1$ then ρ_{i+2} is again selected by $\mathcal{H}_2'$ and thus again $d°P^{(i+2)} \leqslant d_{i+2}-1$. These are the first steps of an induction proof along the idea that if ρ_i is selected by $\mathcal{H}_2'$ then $P^{(i)}$ is formed from $P^{(i-1)}$ and $P^{(\rho_i)}$ and at least one of ρ_{i-1} or ρ_{ρ_i} was selected by $\mathcal{H}_2'$.

The induction hypothesis then proves the lemma.

□

$\mathcal{H}_2'$ is used as long as $N'_{j-1} < 2d_{j-1} - k - 1$ or equivalently $d_j = d_{j-1}$. Thus during a "$\mathcal{H}_2'$ session" we see that the degree remains constant: say d. By the lemma we know that all the corresponding $R^{(j)}$ are in {d-1/d-k-1} . From the moment that $N'_i \geqslant 2d-k-1$ we see that it must be the padé approximant of {d-1/d-k-1}. So it is an inner diagonal element of T_k. From then on we look for an outer diagonal element of T_k, using $\mathcal{H}_1$. Owing to corollary 3.2) we will find all the elements.

□

Proposition 3

If the padé table is normal then for algorithm A3 we have in addition to the statements of theorem 2 that

$$d_i = d_{i-1} + 1 \text{ and then } N'_i = N_i = 2d_i - |k| - 1 \; ; \; d_{i+1} = d_i$$

or

$$d_i = d_{i-1} \quad \text{and then } N'_i = N_i = 2d_i - |k| \; ; \; d_{i+1} = d_i + 1$$

and always $\rho_i = i-2$ and $\mu_i = 1$ (except μ_2 which is $|k|+1$).
$\mathcal{H}_1$ and $\mathcal{H}_2'$ are used alternatively.

Proof :

The proof is trivial and is left to the reader.

□

Proposition 3 implies that for all i we have

$$S^{(i)} = S^{(i-1)} - r_{i-1}\, r_{i-2}^{-1}\, z\, S^{(i-2)}$$

This is the recursion of Watson's and Thacher's algorithm.

Additional remarks

1. If we drop the subscript i everywhere and if we have that

$$R = \frac{P}{Q} \quad \text{and} \quad Q = \sum_{\ell=0}^{d} b_\ell z^\ell \qquad (b_0 = 1)$$

then we can compute N' and r by the following algorithm:

1. K = N
2. compute

$$s_K = \sum_{\ell=0}^{d} b_\ell c_{K-\ell}$$

3. if $s_K = 0$ then K = K+1 and go to 2.
   else r = s_K and N' = K.

If there is no finite K for which $s_K \neq 0$ then f(z) = R

2. If we only want to compute [L/M] for certain prespecified values of L and M then we need not compute all the numerators of the elements preceding [L/M] in the staircase T_k. It is sufficient to use only the recursion for $Q^{(i)}$ and find

$$P^{[L/M]} = \sum_{\ell=0}^{L} a_\ell^{[L/M]} z^\ell$$

from

$$a_\ell^{[L/M]} = \sum_{j=0}^{M} b_j^{[L/M]} c_{\ell-j} \qquad \ell = 0,1,\ldots,L$$

3. Another possibility inspired by the Thacher algorithm would be to compute only F_i using the recursions for $S^{(i)} = Z^{(i)}$

$$Z^{(i)} = Z^{(i-1)} - F_i z^{\mu_i} Z^{(\rho_i)}$$

This however would require for every i an updating of "infinitely" many coefficients of $Z^{(i)}$. This difficulty we can overcome as follows: If we set

$$Z^{(i)} = \sum_{j=i}^{\infty} \alpha_{i,j} z^{j+N_i'}$$

then we have to compute the infinite upper triangular part of the matrix $(\alpha_{i,j})$, every row representing a $Z^{(i)}$. Instead of

doing this row by row, we could do this column by column. This gives rise to some additional housekeeping we will not explain at this place. Another derivation of this Thacher- like version can be found in [2].

4. To put things together we restate algorithm A3 in an algol like notation such that it is fit for implementation.
The problem is : given $\{ c_0, c_1, c_2, \ldots, c_T \}$
find the padé approximant $R_{L_M} \in T_k \quad (M \leq T)$.
Algorithm A3:

```
if k ≤ -0 then { compute Q^[0/k] } ~;
comment initialization ;
  stop = false;found = false;outer = true;
  d̃ = d̄ = 0; Ñ = N̄ = 0;
  if k ≥ +0 then begin r_∅ = -1;{Q̃ = Q̄ = 0 } end
            else begin r_∅ = c_0;{Q̃ = Q̄ = 1 } end ;
  d = N = |k|;
  if k ≥ +0 then {Q = 1} else {Q = Q^[0/-k] };
comment : start computing the staircase ;
for i = 1,i+1 while (stop ∧ found = false) do
  begin comment : determine N and r;
    s=0;
    for K = N,K+1 while (s = 0 ∧ K ≤ T) do {s = Σ_{ℓ=0}^{d} b_ℓ c_{K-ℓ}};
    N = K-1; r = s;
  comment :~
          Q contains Q^(ρ_i)  if ρ_i was selected by 𝓗_1
          Q̄ contains Q^(ρ_i)  if ρ_i was selected by 𝓗'_2
          QT is auxiliary space;
  if outer = true then begin if 2d < N + |k| +1
    then
      begin comment :an outer diagonal element is found; outer=false;
        if N ≥ M
          then begin found=true;{Q = Q-Q̃×z↑(N-Ñ)×r/r̃};end
          else begin {QT = Q};{Q = Q-Q̃×z↑(N-Ñ)×r/r̃};Ñ = N;r̃ = r end
      end                                          {Q̃ = QT};
    else
      if r = 0 then stop = true else {Q = Q-Q̃×z↑(N-Ñ)×r/r̃}
                    end
```

```
                else
  begin
    if 2d < N + |k| +1
      then
        begin comment: an inner diagonal element is found;outer=true;
          if N ≥ M then begin found = true;{Q = Q-Q̄×z↑(N-N̄)×r/r̄} end
                   else begin {QT=Q};{Q=Q-Q̄×z↑(N-N̄)×r/r̄};{Q̄=QT};
                                N̄ = N; r̄ = r
                        end
        end
      else
       if r = 0 then stop = true else {Q = Q-Q̄×z↑(N-N̄)×r/r̄}
  end if;
end for;
```

5.Examples

Take

$$f(z) = \frac{1 - z + z^3}{(1 - z)^2} = 1 + z + z^2 + 2z^3 + 3z^4 + \ldots + nz^{n+1} + \ldots$$

The padé table can be found in appendix 1.

Algorithm A with $\mathcal{H} = \mathcal{H}_1$ would give for T_{-0}:

(the outer diagonal elements of T_k are marked with 0)

i	0	1	2	3	4	5
$R^{(i)}$	$\frac{0}{1}$	$\frac{1}{1}$	$\frac{1}{1-z}$	$\frac{1}{1-z-z^3}$	$\frac{1+z^3}{1-z-z^4}$	$\frac{1-z+z^3}{1-2z+z^2}$
N_i'	0	1	3	6	7	∞
d_i	0	1	1	3	4	4
r_i	1	1	1	-1	-1	0
diagonal	0		0	0		0

criterium $\mathcal{H}_2$ would give

i	0	1	2	3	4	5
$R^{(i)}$	$\frac{0}{1}$	$\frac{1}{1}$	$\frac{1}{1-z}$	$\frac{1-z^2}{1-z-z^2}$	$\frac{1}{1-z-z^3}$	$\frac{1-z+z^3}{1-2z+z^2}$
N_i'	0	1	3	5	6	∞
d_i	0	1	1	3	3	4
r_i	1	1	1	-1	-1	0
diagonal	0		0		0	0

criterium $\mathcal{H}_3$ would give the same as $\mathcal{H}_2$ (for this example)

$\mathcal{H}_i$			$\mathcal{H}_1$	$\mathcal{H}_2'$	$\mathcal{H}_1$	$\mathcal{H}_2'$
diagonal	0	I	0	I	0	I

That we can not replace $\mathcal{H}_2'$ by $\mathcal{H}_2$ in algorithm A3 may be shown by the following example:

$$f(z) = \frac{1+z^4}{1-z^5} = 1 + z^4 + z^5 + z^9 + z^{10} + 2z^{15} + 2z^{20} + 2z^{25} + \ldots$$

The padé table has the form:

L \ M	0	1	2	3	4	5	6	7	8	...
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
:										

T_{+0} computation by A3 gives:

i	0	1	2	3	4	5
$R^{(i)}$	$\frac{1}{0}$	$\frac{1}{1}$	$\frac{1+z^4}{1}$	$\frac{1-z+z^4}{1-z}$	$\frac{1-z+z^2+z^4}{1-z+z^2}$	$\frac{1-z+z^2-z^3+z^4}{1-z+z^2-z^3}$
N_i'	0	4	5	6	7	8
d_i	0	1	4	4	4	4
r_i	-1	1	1	-1	1	-1
$\mathcal{H}_i$			$\mathcal{H}_1$	$\mathcal{H}_1$	$\mathcal{H}_1$	$\mathcal{H}_1$
diag	0	I				0

i	6	7	8
$R^{(i)}$	$\frac{1-z+z^2-z^3+2z^4}{1+z+z^2-z^3+z^4}$	$\frac{1+z-z^2+z^3+2z^5}{1+z-z^2+z^3-z^4}$	$\frac{1+z^4}{1-z^5}$
N_i'	9	10	∞
d_i	5	5	6
r_i	2	2	0
$\mathcal{H}_i$	$\mathcal{H}_2'$	$\mathcal{H}_1$	$\mathcal{H}_2'$
	I	O	I

If $\mathcal{H}_2$ would have been used instead of $\mathcal{H}_2'$ then $R^{(6)}$ would have been

$$R^{(6)} = \frac{1+z^4+z^5}{1} \qquad \text{and } R^{(7)} = \frac{1+z-z^2+z^4+2z^5}{1+z-z^2+z^3-z^4}$$

$R^{(7)}$ is on the outer diagonal so the inner diagonal element is jumped over.

Acknowledgment

We thank prof. P.De Wilde who introduced the Berlekamp-Massey algorithm to us and prof. L.Wuytack who brought the papers of Bussonnais and Cordellier to our attention.

References

[1] C.Brezinski:
Computation of padé approximants and continued fractions.
Journ. of Comp. and Appl. Math. 2(2) p.113-123,1976

[2] A.Bultheel
Division algorithms for continued fractions and the padé table.
Katholieke Universiteit Leuven, Applied Math. and Progr. Division, aug.1978.

[3] D.Bussonnais:
"Tous" les algorithmes de calcul par recurrence des approximants de padé d'une serie. Construction de fractions continues correspondantes.
Séminaire d'Analyse Numérique,n° 293,Grenoble 1978

[4] G.Claessens:
A new look at the padé table and the different methods for computing its elements.
Journ. of Comp.and Appl.Math. 1(3) p 141-152,1975

[5] G.Claessens -L.Wuytack:
On the computation of non normal padé approximants
Universitaire instelling Antwerpen (UIA),Rept. 77-41,nov. 1977.

[6] B.Cordellier:
Deux algorithmes de calcul récursif des éléments d'une table de padé non normale.
Presented at the conference on padé approximation,Lille,1978

[7] L.S.De Jong:
Numerical aspects of realization algorithms in linear systems theory.
Thesis. T.H.Eindhoven, The Netherlands, 1975.

[8] W.B.Gragg:
The padé table and its relation to certain algorithms of numerical analysis.
SIAM Review 14,p.1-62, 1972.

[9] J.L.Massey:
Shift-register synthesis and BCH decoding.
IEEE Transactions on Information Theory,IT-15(1),p.122-127,1969.

[10] P.J.S.Watson:
Algorithms for differentiation and integration
in: "Padé approximants and their applications" P.R.Graves-Morris,ed.,Ac.Press,London,p.93-98,1973.

Appendix 1

Example of a non normal padé table (Gragg)

$$\frac{z^2(z^3-z+1)}{(z-1)^2} = z^2+z^3+z^4+2z^5+\ldots+nz^{n-1}+\ldots$$

L \ M	0	1	2	3	4	5	...
0	$\frac{0}{1}$	$\frac{0}{1}$	nill	nill	nill	nill	...
1	$\frac{0}{1}$	nill	nill	nill	nill	nill	
2	$\frac{z^2}{1}$	$\frac{z^2}{1-z}$	$\frac{z^2}{1-z-z^3}$				
3	$\frac{z^2(1+z)}{1}$	nill				nill	
4	$\frac{z^2(1+z+z^2)}{1}$				nill	nill	
5			$\frac{z^2(1-z+z^3)}{(z-1)^2}$				

$$[4/1] = \frac{z^2(1-z-z^2)}{1-2z} \quad ; \quad [4/2] = \frac{z^2(1-z^2)}{1-z-z^2} \quad ; \quad [5/0] = \frac{z^2(1+z^3+z^4+2z^5)}{1}$$

$$[5/1] = \frac{1 - 3/2\ z - 3/2\ z^2 - 1/2\ z^3 + z^4 + 2\ z^5}{1 - 3/2\ z}$$

The padé table of $\frac{z^3-z+1}{(z-1)^2}$ is obtained by deleting the first two rows and the factor z^2 in the remaining table.

Appendix 2

Proposition 4

The order of {L/M} is the order of the indicator of [L/M]

Proof:

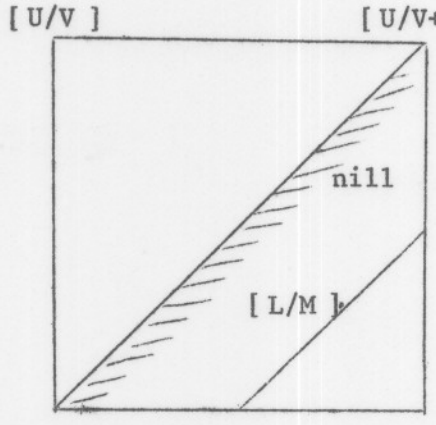
That the order of {L/M} is not less than the order of the indicator of [L/M] is trivial since this indicator is one of the elements of {L/M}.

That equality holds is proved if $(L/M)_K$ = nill for all K > N.

Suppose there is a non nill element $(L/M)_K \in \{L/M\}$ having an order K > N where N is the order of the indicator.

If [L/M] ≠ nill then [L/M] is itself the indicator. It follows from the definition of $(L/M)_K$ and of [L/M] that $(L/M)_K$ = [L/M] because [L/M] is uniquely defined. Thus $(L/M)_K$ = nill which is a contradiction.

If [L/M] = nill, let [U/V] be its indicator. We have K > N=U+V+n+1



It is known [8,p.13] that a solution of the determining conditions of $(L/M)_K$ does not exist. Only if we do not require that it should be a reduced and normalized fraction, then it has a solution of the form

$$\frac{z^{\omega} P^{U}(z)}{z^{\omega} Q^{V}(z)}$$ for some ω. Thus

$(L/M)_K$ can only be nill for K > U+V+n+1. This proves the proposition. □