to appear in continued fractions and Padé approximants.
C. Brezinski (ed.) North Holland, 1990

A matrix Euclidean algorithm and the matrix minimal Padé approximation problem

Adhemar Bultheel and Marc Van Barel
Department of Computer Science, K.U.Leuven

ABSTRACT

We formulate a minimal Padé approximation problem and generalize it for rectangular matrix series. The scalar solutions can be computed by the Euclidean algorithm and we give its generalization to compute the matrix approximants. The algorithm is normalized so that it generates unique matrix minimal Padé approximants and a unique continued fraction representation for them.

0. Introduction

The problem of Padé approximation is well known for scalar series $F(z)$. If we allow the series to be $p\times m$ matrix series, it is not trivial to extend the theory. To give an idea of the problems that may arise, we note the following.

A scalar rational form $N(z)M(z)^{-1}$ can be simply normalized by requiring e.g. that $M(z)$ is monic. If however $N(z) \in \mathbb{C}^{p\times m}[z]$ and $M(z) \in \mathbb{C}^{m\times m}[z]$, the normalization is not simple. We have chosen for the simplest possible (in a certain sense) representation of $M(z)$, i.e. for a canonical form of $M(z)$.

Another problem which arises in the matrix case is the degree of numerator and denominator which is needed in the definition of Padé approximant. In the scalar case, there is only one degree for numerator and one for the denominator. For the matrix case, one could consider all possible kinds of degree information since each entry of the numerator and the denominator matrix is a scalar polynomial with a specific degree. This seems to be too detailed a specification of the degrees in the general matrix case, although it works for vector approximants. The other extreme is to consider numerator and denominator as polynomials with matrix coefficients. In that case we associate again one degree with the numerator and one degree with the denominator. This seems to work partially for the square matrix case when we have a normal Padé table. However, a matrix coefficient can be singular without being zero and these situations of the non normal case ask for more detailed degree information. We shall here consider row degrees for the matrices. This is something in between the two previous extremes. Thus we should write $N(z) \in (\mathbb{C}^m[z])^p$ rather than $N(z) \in \mathbb{C}^{p\times m}[z]$, but we shall not do so. It is then a logical consequence to consider also the orders of approximation rowwise.

Our approach will be via the Euclidean algorithm. It is known that in the scalar case, it can be used to generate Padé approximants found along an anti diagonal of the Padé table. I.e. with at least a certain fixed order of approximation and with increasing denominator degree and decreasing numerator degree. The Padé approximants generated by the Euclidean algorithm are solutions of what we shall call a minimal Padé approximation problem (a term akin to minimal

partial realization). Suppose one raises the question : which is the rational approximant with minimal denominator degree that gives a certain order of approximation in a Padé sense. Of course if there is no restriction on the numerator degree, this minimal Padé approximation problem always has a trivial solution : take the denominator equal to 1 and the numerator equal to the required partial sum. Thus a minimal Padé approximation problem can only have sense if such a restriction is given. The restriction we have chosen here will be formulated in problem MPA(δ). A matrix generalization of this problem is then given as well as a matrix generalization of the Euclidean algorithm which will generate its solutions. We shall call these solutions matrix minimal Padé approximants.

We shall also give a continued fraction representation of the minimal Padé approximants generated by this algorithm. This representation can be easily set up by the quantities appearing in the algorithm. To be sure we get a normalized approximant we shall impose some normalizing conditions in the Euclidean algorithm. Moreover some extra conditions are introduced to make the complete Euclidean algorithm unique, hence also the continued fraction representations.

Matrix Padé approximants and more generally Padé approximants in a noncommutative algebra have been considered by many authors before us. Draux [8] has commented on about 300 references on this topic. The vector case i.e. ($m=1$) has been considered by de Bruin [6] and Graves-Morris [10]. More recently matrix Padé approximants were considered, under certain restrictions in [20,21]. It is however in the literature on the closely related minimal partial realization problem for multi-input-multi-output systems that the fine structure of the row and/or column degree information of numerator and denominator are fully exploited for the general case. The main references here are [2] and [7]. In this field Fuhrmann [9] describes a matrix Euclidean algorithm and a matrix continued fraction. His results do not seem to be very successful when one tries to apply them to the Padé problem. While the authors were in the process of preparing the series of reports [13-17] on algorithms for minimal partial realization in the matrix case, a paper of Antoulas [3] appeared which solves a lot of problems in this area. He also gives a matrix Euclidean algorithm and a continued fraction associated with it. While Antoulas' paper is more on a theoretical side, our reports emphasise the algorithmic aspects. Inspired by the work of Antoulas we gave in [17] a canonical version of the algorithm which turns out to be suitable for the computation of matrix Padé approximants as they are defined in this paper.

Notation:

- PA: Padé approximant
- gcd : greatest common divisor.
- gcld : greatest common left divisor.
- $\underline{\mu}(\delta) = \{\mu_1(\delta), \ldots, \mu_m(\delta)\}$: Kronecker indices of $\{F_\omega, \ldots, F_{-\delta+1}\}$
- $\underline{\lambda}(\delta) = \{\lambda_1(\delta), \ldots, \lambda_p(\delta)\}$: dual Kronecker indices of $\{F_\omega, \ldots, F_{-\delta+1}\}$
- ∂M : degree of polynomial $M(z)$.

- $F(z) = O_-(z^\omega)$ iff $F(z) = F_\omega z^\omega + F_{\omega-1} z^{\omega-1} + \cdots$
- $F(z) = O_+(z^\omega)$ iff $F(z) = F_\omega z^\omega + F_{\omega+1} z^{\omega+1} + \cdots$
- $F(z) = O_-(z^{\gamma_{ij}})$ iff $F(z) = [f_{ij}(z)]$ with $f_{ij}(z) = O_-(z^{\gamma_{ij}})$.
- $|\underline{\mu}| = \sum \mu_i$
- I_p : $p \times p$ identity matrix.
- M_{hc} : matrix of leading column coefficients.
- M_{hr} : matrix of leading row coefficients.

1. The scalar minimal Padé problem

The scalar Padé approximation problem is well known [4]. There exist several ways to characterize a Padé approximant (PA). The Baker definition goes as follows. Given a formal power series $F(z) = \sum_{k=0}^{\infty} F_k z^k$, with $F_k \in \mathbb{C}$, a $[\alpha/\beta]$ PA is defined as the rational form $N(z)M(z)^{-1}$ satisfying

1. $N(z)$ has degree $\nu \leqslant \alpha$
2. $M(z)$ has degree $\mu \leqslant \beta$
3. $F(z) - N(z)M(z)^{-1} = O_+(z^{\gamma+1})$ with $\gamma \geqslant \alpha + \beta \geqslant \nu + \mu$
4. $M(z)$ and $N(z)$ coprime

The table containing at the intersection of row α and column β the $[\alpha/\beta]$ approximant is called the Padé table.

Not all the entries of the Padé table have to exist. In general, if $F_0 \neq 0$, the table consists of square blocks of finite dimension. If $F(z)$ is the expansion of a rational function $P(z)Q(z)^{-1}$ ($P(z)$ and $Q(z)$ coprime) an infinite block with left top element $P(z)Q(z)^{-1}$ can occur. In a finite block, all elements on and above its main anti diagonal are equal to the simplest element in that block, viz. the left top element. Below this anti diagonal, the PAs do not exist according to the previous definition.

To ensure existence of all $[\alpha/\beta]$ approximants we use the older Frobenius type definition of a PA. These approximants, as defined below, do exist for every α and β and it are these approximants that are computed by the Euclidean algorithm.

Definition PA: The rational form $N(z)M(z)^{-1}$ is a $[\alpha/\beta]$ PA of $F(z)$ iff

1. $N(z) = O_-(z^\alpha)$ (1.1)
2. $M(z) = O_-(z^\beta)$ (1.2)
3. $F(z)M(z) - N(z) = O_+(z^{\gamma+1})$ (1.3)
4. $\gamma \geqslant \alpha + \beta$ (1.4)

5. $M \not\equiv 0$ (1.5)

6. ∂M as low as possible. (1.6)

The $[\alpha/\beta]$ approximant can be normalized by imposing that $M(z)$ should be monic.

The Frobenius definition [4] is the above definition without condition 6. This allows a complete equivalence class of PAs for entries in a singular block of the Padé table. A single one could be selected by taking the reduced form, i.e. the rational function defined by it. Then all the PAs in a square block are equal to the left top element. Also those below the main anti diagonal.
Since the Euclidean algorithm does compute entries of the Padé table that are below the anti diagonal of a square block in an unreduced form, we define the PAs by conditions 1-5 and make them as simple as possible by imposing condition 6. What we get is illustrated by the example below.
To illustrate the difference between both definitions, consider the example $F(z) = 1 + z^4 + z^5 + z^9 + \cdots$. Its Padé table has a 4 by 4 block in its left top corner. It can be filled in as in the tables below.

	0	1	2	3	β
0	1/1	1/1	1/1	1/1	
1	1/1	1/1	1/1		
2	1/1	1/1			
3	1/1				
α					

Baker definition

	0	1	2	3	β
0	1/1	1/1	1/1	1/1	
1	1/1	1/1	1/1	z/z	
2	1/1	1/1	z/z	z^2/z^2	
3	1/1	z/z	z^2/z^2	z^3/z^3	
α					

Definition PA

The Frobenius definition has 1/1 all over the square. Note that all definitions are equivalent iff $M(0) \not\equiv 0$.

We shall now introduce a family of minimal Padé approximation problems, MPA(δ), parametrized by an integer δ. The solutions will be PAs that are computed by the Euclidean algorithm.
Suppose we are given the order ω, and ask which is the simplest PA with numerator $N(z)$ and denominator $M(z)$ that satisfies (1.3), then it will depend on the relative importance of numerator degree and denominator degree which PA we shall select. Such a relative importance could e.g. be imposed by requiring that $N(z)M(z)^{-1} = O_-(z^{-\delta})$ for some δ ($|\delta| \leq \omega$). So we consider the following problem for fixed $F(z)$ and ω :

Problem MPA(δ) : Given a power series $F(z)$, a nonnegative integer ω and an integer $\delta \in [-\omega, \omega]$, find nontrivial polynomials $M(z)$ and $N(z)$ such that

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (1.7)

2. $N(z)M(z)^{-1} = O_-(z^{-\delta})$ (1.8)

3. $\mu = \partial M(z)$ as low as possible. (1.9)

If there is still more than 1 solution for (1.7-9), then select the one for which

4. $\nu = \partial N(z)$ as low as possible. (1.10)

The solution for this problem always exists and it is unique up to a constant multiple in numerator and denominator. Hence we can pick unique representations by choosing e.g. a monic denominator. If $M(z)$ has to satisfy (1.7) with $\partial M = \mu$ and $\partial N = \nu$, then the coefficients of $M(z)$ should satisfy the Hankel system

$$\sum_{l=0}^{\mu} F_{\omega-i-l} M_l = 0 \quad , \quad i = 0,1,\ldots,\omega-\nu-1 \tag{1.11}$$

where $M(z) = \sum_0^{\mu} M_j z^j$.

To satisfy (1.8), we must have $\nu \leqslant \mu-\delta$. Thus to satisfy (1.7-9), we try first $\mu = 0$, then $\mu = 1$ etc. I.e. we successively try to find nontrivial solutions for the Hankel systems ($F_{-k} = 0$ for $k > 0$)

$$\begin{bmatrix} F_\omega & \cdots & F_{\omega-\mu} \\ \vdots & & \vdots \\ F_{\mu-\delta+1} & \cdots & F_{-\delta+1} \end{bmatrix} \begin{bmatrix} M_0 \\ \vdots \\ M_\mu \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad , \quad M_\mu \neq 0 \tag{1.12}$$

for $\mu = 0,1,\ldots,\omega+\delta$. For $\mu = \omega+\delta$ there is no system (1.12) so that $M(z)$ can be chosen to be an arbitrary polynomial of degree $\mu = \omega+\delta$ and also for $\delta = -\omega$, the system (1.12) does not exist. $N(z) = \sum_0^{\omega} F_k z^k$ and $M(z) = 1$ is then always a solution. The first μ for which the last column of the Hankel matrix of the systems (1.12) is linearly dependent on the previous columns will give the minimal μ for which a solution of problem MPA(δ) exists. We represent this minimal μ as $\mu(\delta)$. As we shall see later, there will always exist a solution of problem MPA(δ) with $\nu \leqslant \omega - \mu(\delta)$ (see (3.9)), so that $\nu + \mu(\delta) \leqslant \omega$. It may happen that $N(z) = 0$, in which case we set $\nu = -\infty$. To avoid exceptional rules for the case $N(z) = 0$, we add an artificial row -1 to the Padé table with entries $[-1/k] = 0/z^k$. With this extension we can then say that a solution for problem MPA(δ) will be a $[\omega-\mu(\delta)/\mu(\delta)]$ PA.

The indices $\mu(\delta)$ are the scalar version of the Kronecker and dual Kronecker indices for the sequence $F_\omega, F_{\omega-1}, \ldots, F_{-\delta+1}$ in the matrix case (see the next section). These play an essential role in the theory of the Euclidean algorithm. Note that in the scalar case the Kronecker and dual Kronecker indices coincide and that $\mu(-\omega) = 0$ and $\mu(\delta) \leqslant \omega+\delta$. For a normal Padé table one can verify that we will find that $\mu(\delta) = (\omega+\delta)/2$ and $\nu = (\omega-\delta)/2$ for $\omega+\delta$ even and $\mu(\delta) = (\omega+\delta+1)/2$ and $\nu = (\omega-\delta-1)/2$ for $\omega+\delta$ odd.

A geometrical interpretation of the problem is simple for a normal Padé table. Draw the upward sloping diagonal $\nu + \mu = \omega$, which cuts off the left top corner of the Padé table. Consider the remaining set S_ω of all $[\nu/\mu]$ PAs with $\nu + \mu \geqslant \omega$. Every element in S_ω will satisfy condition (1.7). Next, draw the downward sloping diagonal $\mu - \nu = \delta$. In a normal Padé table, all the entries above this diagonal will satisfy $\mu - \nu \geqslant \delta$, i.e. they satisfy (1.8). Call this set S_δ. The intersection $S_\omega \cap S_\delta$ is the eastern quadrant defined by the two diagonals. The elements with minimal denominator degree are the leftmost elements in this quadrant close to the intersection of the diagonals. Depending on the values of ω and δ, there is only one or there are two solutions. In the latter case choose the one with minimal numerator degree, i.e. the one that is above the other.

If the Padé table is not normal and has singular blocks, the situation is a bit more complicated. One also has to drop from $S_\omega \cap S_\delta$ those PAs that belong to a block whose main diagonal is

below the diagonal bordering S_δ and then proceed as before.
Note that the elements above the main diagonal of a block will never be selected as a solution for problem MPA(δ).
For our previous example, the entries of the block that solve problem MPA(δ) for some values of ω and δ are given below.

	0	1	2	3	β
0	1/1				
1	1/1	1/1			
2	1/1	1/1	z/z		
3	1/1	z/z	z^2/z^2	z^3/z^3	
α					

Solutions of problem MPA(δ).

As you can see, not every PA will be a solution of problem MPA(δ), but if some PA (above the diagonal of the block) is not a solution, it will be equal to some other PA (on or below the diagonal) which is a solution.

It is possible to describe the Padé table in terms of ω and δ instead of α and β. While α indicates a row and β a column, ω refers to an anti diagonal and δ to a main diagonal and all entries, block structure etc. can be described by the diagonal grid $(\omega,\delta), \omega = 0,\ldots,\infty$; $\delta = -\omega,\ldots,\omega$. Thus instead of defining $[\alpha/\beta]$ PAs, we could define $[\omega,\delta]$ PAs. To distinguish it from the classical terminology, we shall call the latter a $[\omega,\delta]$ MPA and its definition is the properly normalized solution of problem MPA(δ) for some given ω. The $[\omega,\delta]$ table will look in the normal case as the usual Padé table rotated over 45 degrees but with every element (except those with $\beta = 0$) duplicated. The same thing will happen if there are blocks but the entries of the (ω,δ) table corresponding to the upper triangular part of the previous picture, i.e., the right triangular part of the (ω,δ) block, will be filled up with the entries immediately outside the (α,β) block on the same anti-diagonal in the upward sloping direction. This explains why we needed the artificial line −1. To make no exception for $\beta = 0$ we could introduce an extra δ-value viz. $\delta = -\omega-1$ and set the solution of MPA(δ) for this value equal to the solution for MPA($-\omega$). The structure can be seen in the figures of section 3.

2. The matrix minimal Padé problem

We shall now give the natural matrix generalization of the theory of the previous section for the case that the $F_k \in \mathbb{C}^{p\times m}$. Whether we can call its solutions Padé approximants will depend on how we extend the definition of Padé approximant. For the time being, we shall call it a solution of the matrix minimal Padé problem (problem MPA(δ)) and call it an $[\omega,\delta]$ MPA. A justification for calling it a Padé approximant will be given at a later stage.

Suppose we have a formal power series

$$F(z) = \sum_{k=0}^{\infty} F_k z^k \quad \text{with} \quad F_k \in \mathbb{C}^{p\times m}$$

In the extension of problem MPA(δ) we now consider polynomial matrices $N(z) \in \mathbb{C}^{p\times m}[z]$ and $M(z) \in \mathbb{C}^{m\times m}[z]$. Again we can look for the simplest rational form which approximates up to a

given order and impose a relative importance to numerator and denominator degree by requiring $N(z)M(z)^{-1} = O_-(z^{-\delta})$. Therefore we consider problem MPA(δ) which is almost as in the scalar case.

Problem MPA(δ) : Given $F(z)$ as above, some nonnegative integer ω and an integer $\delta \in [-\omega,\omega]$, find polynomials $M(z)$ and $N(z)$ as above such that

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (2.1)
2. $N(z)M(z)^{-1} = O_-(z^{-\delta})$ (2.2)
3. Suppose the denominator has column degrees $\underline{\mu} = \{\mu_1, \ldots, \mu_m\}$. Then make

$$|\underline{\mu}| = \sum_{j=1}^{m} \mu_j \text{ as low as possible.} \tag{2.3}$$

If there is more than one solution satisfying (2.1-3), then choose the one such that

4. some uniqueness condition is satisfied. (2.4*)

To specify condition (2.4*), we need the notion of Kronecker and dual Kronecker indices first. These appear in a natural way when we analyse problem MPA(δ) somewhat deeper.
Let us denote the columns of $M(z)$ by

$$m_j(z) = \sum_{i=0}^{\mu_j} m_{ij} z^i \quad , \quad m_{ij} \in \mathbb{C}^m.$$

First note that condition (2.3) implicitly requires that $M(z)$ has a regular matrix of leading column coefficients. If it were not so, it would be possible by taking polynomial linear combinations of the columns of $M(z)$ to reduce the degree of some of the columns and still satisfy (2.1-2). It would then not be a minimal degree solution. Hence

$$M_{hc} = [m_{\mu_1,1} \cdots m_{\mu_m,m}]$$

should be nonsingular. Such a polynomial matrix is called *column reduced* and the degree of its determinant is equal to

$$\partial \det M(z) = \sum_{j=1}^{m} \mu_j = |\underline{\mu}|.$$

Now let us try to find the minimal column degrees $\underline{\mu}(\delta)$ required to solve problem MPA(δ). The columns of $M(z)$ have to satisfy some block Hankel systems just like (1.11)

$$\sum_{i=0}^{\mu_j} F_{\omega-i-t} m_{ij} = 0 \quad , \quad t = 0,1,2,\ldots,\omega-\eta_j-1 \quad , \quad j = 1,2,\ldots,m$$

where the η_j are defined by $N(z) = O_-(z^{\eta_j})$. Moreover we have to make M_{hc} nonsingular. Since condition (2.2) requires $\eta_j \leqslant \mu_j - \delta$ if $M(z)$ is column reduced, we can consider as in the scalar case the block Hankel systems (we suppose again that $F_{-k} = 0$ for $k > 0$)

$$\begin{bmatrix} F_\omega & \cdots & F_{\omega-\mu_j} \\ \vdots & & \vdots \\ F_{\mu_j-\delta+1} & \cdots & F_{-\delta+1} \end{bmatrix} \begin{bmatrix} m_{0j} \\ \vdots \\ m_{\mu_j,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.5}$$

for $\mu_j = 0,1,\ldots,\omega+\delta$ as low as possible and M_{hc} regular. This problem is well known in the

theory of block Hankel matrices and the minimal μ_js are known as the Kronecker indices of the sequence $F_\omega, F_{\omega-1}, \ldots, F_{-\delta+1}$. We shall define them through a practical scheme that can be used to compute them. It has the advantage that it will then be fairly simple to see that they are indeed the minimal degrees needed in problem MPA(δ).

Definition (Kronecker indices) : Let $F_\omega, \ldots, F_{-\delta+1}$, $-\omega \leqslant \delta \leqslant \omega$ be given, all matrices from $\mathbb{C}^{p\times m}$. For $j = 1, \ldots, m$ call h_j the j-th column of the matrix

$$[F_\omega^T, F_{\omega-1}^T, \ldots, F_{-\delta+1}^T]^T$$

and place h_j at entry $(0,j)$ of the column table. These columns have length $(\omega+\delta)p$. Next call h_{m+j}, $j = 1, \ldots, m$ the j-th column of the matrix

$$[F_{\omega-1}^T, F_{\omega-1}^T, \ldots, F_{-\delta+1}^T]^T$$

and place h_{m+j} at entry $(1,j)$ of the column table. These columns have length $(\omega+\delta-1)p$. Continue like that, each time dropping the first block entry of the sequence until the m columns $h_{(\omega+\delta-1)m+j}$ of $F_{-\delta+1}$ are placed at positions $(\omega+\delta-1, j)$ of the column table. From this table we construct the crate table which defines linear (in)dependencies of these columns in the following way: If column $h_{im+j} \in \mathbb{C}^{(\omega+\delta-i)p}$ is linear independent of all preceding columns truncated at the same length, we place a × at entry (i,j), otherwise we place a ●. If we run out of entries in the column table, then just place a ● at the corresponding place in the crate table.

	1	2	⋯	m
0	h_1	h_2	⋯	h_m
1	h_{m+1}	h_{m+2}	⋯	h_{2m}
2	h_{2m+1}	h_{2m+2}	⋯	h_{3m}
3	h_{3m+1}	h_{3m+2}	⋯	h_{4m}
⋮	⋮	⋮		⋮

column table

	1	2	⋯	m
0	×	×	⋯	×
1	●	×	⋯	×
2	●	×	⋯	●
3	●	●	⋯	●
⋮	⋮	⋮		⋮

crate table

Let $\mu_j(\delta)$ be the row index of the first dot appearing in column j of the crate table. Then we define $\underline{\mu}(\delta) = \{\mu_1(\delta), \ldots, \mu_m(\delta)\}$ as the Kronecker indices of the sequence $F_\omega, \ldots, F_{-\delta+1}$.

Note that it is a property of the block Hankel structure that if in a certain column of the crate table a dot appears, then all the following entries of that column will be dots too.
Note also that the entries of the column table are the columns of a block Hankel matrix H where only the upper left triangular block part is specified. As a matter of fact, if we can mark a dot in the crate table, this coincides with the fact that a system of the form (2.5) has a solution with a leading coefficient $m_{\mu_j,j}$ which is linearly independent from the other ones. It is obvious from these observations that we have proved the following theorem :

Theorem 1 : *The minimal column degrees for $M(z)$ needed to solve problem MPA(δ) will be the Kronecker indices $\underline{\mu}(\delta)$ of the sequence $F_\omega, \ldots, F_{-\delta+1}$.*

□

The dual Kronecker indices $\lambda_i(\delta)$, $i = 1,2, \ldots, p$ are defined in a similar way, now working with the rows of

$$[F_\omega, F_{\omega-1}, \ldots, F_{-\delta+1}].$$

Moreover the number of independent columns must be equal to the number of independent

rows and equal to the rank of the matrix H. Thus

$$|\underline{\lambda}(\delta)| = \sum_{i=1}^{p} \lambda_i(\delta) = |\underline{\mu}(\delta)| = \sum_{j=1}^{m} \mu_j(\delta) = r$$

where r is the rank of the block Hankel matrix H where the unspecified elements are supposed to be chosen so that they do not influence the (in)dependencies of rows and columns.
Even if we satisfy the minimality condition for the μ_j, there may still be many solutions to problem MPA(δ). However, up to a normalization, there is only one which satisfies

$$N(z)M(z)^{-1} = O_{-}(z^{\omega-\lambda_i(\delta)-\mu_j(\delta)}). \tag{2.4}$$

This is the condition to replace (2.4*). In the scalar case this reduces to $N(z)M(z)^{-1} = O_{-}(z^{\omega-2\mu})$, and this is equivalent with condition (1.10). That there is only one solution for problem MPA(δ) satisfying (2.1-4) can be derived from system theoretic results found e.g. in [5]. We shall not do so and deliver proofs independent of system theory. The only freedom that is left now to fix $N(z)$ and $M(z)$ uniquely is a normalization. In the scalar case this freedom was the multiplication with a nonzero constant. In the matrix case, a transformation which doesn't change the rational form $N(z)M(z)^{-1}$ and which doesn't change the degree of the determinant of $M(z)$ is a multiplication of $N(z)$ and $M(z)$ from the right with a *unimodular* matrix. This is a polynomial matrix with a nonzero constant determinant. The inverse of a unimodular matrix is again a unimodular matrix. This right unimodular multiplication can be used to make $M(z)$ CURI canonical by which we mean that it is row as well as coumn reduced with leading Column coefficient matrix unit Upper triangular and leading Row coefficient matrix the Identity matrix. Thus

Definition (CURI canonical)

We call a square polynomial matrix $M(z) \in \mathbb{C}^{m\times m}[z]$ CURI canonical with degrees μ_j, $j = 1, \ldots, m$ iff

1. The j-th column of $M(z)$ has degree μ_j, $j = 1, \ldots, m$, (2.6)
2. The i-th row of $M(z)$ has degree μ_i, $i = 1, \ldots, m$, (2.7)
3. The leading column coefficient matrix M_{hc} is unit upper triangular, (2.8)
4. The leading row coefficient matrix M_{hr} is the unit matrix. (2.9)

Thus a CURI canonical polynomial matrix is column as well as row reduced. An algorithm to bring a polynomial matrix in echelon canonical form can be found in [12]. Our definition of CURI canonical is almost the same as the echelon form up to a permutation of the columns, so essentially the same algorithm can be used. It consists of Gaussian type elementary transformations on the columns of $M(z)$ where the multiples of columns added to other columns are now polynomial multiples instead of constant multiples as in the case of a constant matrix. A possible algorithm can be found in appendix E (see Make_M_canonical).
In our case the row and column degrees μ_j will be the Kronecker indices $\mu_j(\delta)$. For an arbitrary matrix $M(z)$, one can always consider the μ_j as Kronecker indices of $F_\omega, \ldots, F_{-\delta+1}$ for some ω and δ where $N(z)M(z)^{-1} = \sum F_k z^k$ with $N(z)$ arbitrary but such that $N(z)$ and $M(z)$ are right coprime.

To summarize what we have found so far, we can formulate the following theorem:

Theorem 2 : *Let* $F(z) = \sum_{k=0}^{\infty} F_k z^k$ *be a formal power series with* $F_k \in \mathbb{C}^{p\times m}$. *Suppose that* ω *is a non negative integer and* δ *an integer from the interval* $[-\omega,\omega]$. *Let* $\underline{\mu}(\delta)$ *be the Kronecker indices of the sequence* $F_\omega, F_{\omega-1}, \ldots, F_{-\delta+1}$. *Then problem MPA(*$\delta$*) has exactly one solution with a CURI canonical denominator with degrees equal to* $\underline{\mu}(\delta)$.

Proof : The proof of the existence is included in the proofs of theorems 4 and 5 of section 4. These proofs are constructive and will include an algorithm to compute the solution of MPA(δ). They are given in appendices B and C. The uniqueness is proved in appendix A.

□

The unique solution described in the previous theorem will be called the $[\omega,\delta]$ MPA of $F(z)$. Now we could try and give the α-β parametrization. This is much more difficult than in the scalar case since ω and δ are two integers while the row degrees of $N(z)$ and $M(z)$ are vectors of integers. So we shall consider a set of row degrees $\underline{\nu} = \{\nu_1, \ldots, \nu_p\}$ for $N(z)$ and a set of row degrees $\underline{\mu} = \{\mu_1, \ldots, \mu_m\}$ for $M(z)$. Since $M(z)$ is CURI canonical, its row degrees are equal to its column degrees. It is both row and column reduced. Our definition of matrix Padé approximant is as follows :

Definition PA: The rational form $N(z)M(z)^{-1}$ is a $[\alpha/\beta]$ PA of $F(z)$ for some non negative integers α and β iff

1. $N(z) \in \mathbb{C}^{p\times m}[z]$ has row degrees $\underline{\nu}$ and $\alpha \geqslant |\underline{\nu}|$ (2.10)
2. $M(z) \in \mathbb{C}^{m\times m}[z]$ is CURI canonical with row degrees $\underline{\mu}$ and $\beta \geqslant |\underline{\mu}|$ (2.11)
3. $F(z)M(z) - N(z) = O_+(z^{\omega_i+1})$ (2.12)
   thus a residual with row orders at least $\underline{\omega} = \{\omega_1, \ldots, \omega_p\}$
4. $\sum_{i=1}^{p} \omega_i \geqslant \sum_{i=1}^{p} \nu_i + \sum_{j=1}^{m} \mu_j$ i.e. $\gamma = |\underline{\omega}| \geqslant \alpha + \beta \geqslant |\underline{\nu}| + |\underline{\mu}|$. (2.13)
5. $\partial \det M = |\underline{\mu}|$ as low as possible (2.14)

Note that in the vector case with $m = 1$ we get a scalar denominator which gives simultaneous Padé approximants with a common denominator. However our definition does not reduce to the classical definition of simultaneous Padé approximants since we only specify an upper bound α for the sum of the numerator degrees. An increase in α can be used to increase any of the numerator degrees. Also the corresponding increase in the order of approximation can be freely obtained from any of the residuals and it need not be the residual of the approximant whose numerator degree has been increased. Still for the vector case, it is simple to check that the number of free parameters is precisely equal to the number of conditions one imposes on the approximant. However since we can distribute these conditions over the different approximants we loose unicity as we had in the scalar case. Consider e.g. the example

$$F(z) = \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix} z + \begin{bmatrix}1\\0\end{bmatrix} z^2 + \cdots .$$

Then $[1\ 0]^T/1$ is the only $[0/0]$ approximant, but there are two $[1/0]$ approximants: $[1+z\ \ 0]^T/1$ and $[1\ \ z]^T/1$. For the $[2/0]$ approximants, there are 3 solutions etc.

However, once we have chosen the normalized solution of problem MPA(δ) with CURI canonical denominator, then it will be a PA in the sense of the above definition. Indeed, for this solution we have

$$N(z) = O_-(z^{\omega-\lambda_i(\delta)-\mu_j(\delta)})M(z) = O_-(z^{\omega-\lambda_i(\delta)}),$$

since by normalization $M(z)$ is row reduced with row degrees $\mu_i(\delta)$. Hence the row degrees ν_i of $N(z)$ are bounded by

$$\nu_i \leqslant \omega-\lambda_i(\delta)$$

so that

$$\sum_{i=1}^{p}\nu_i + \sum_{i=1}^{m}\mu_j \leqslant \sum_{i=1}^{p}(\omega-\lambda_i(\delta)) + \sum_{i=1}^{m}\mu_j(\delta)$$
$$= p\omega - \sum_{i=1}^{p}\lambda_i(\delta) + \sum_{i=1}^{m}\mu_i(\delta) = p\omega = |\underline{\omega}|$$

which means that it is a PA corresponding to our definition. Thus we have proved the following theorem :

Theorem 3 : *Every* $[\omega,\delta]$ *MPA for* $F(z) = \sum_{k=0}^{\infty} F_k z^k$ *is also a* $[|\omega-\underline{\lambda}(\delta)|/|\underline{\mu}(\delta)|]$ *PA where* $\underline{\mu}(\delta)$ *and* $\underline{\lambda}(\delta)$ *are the Kronecker and dual Kronecker indices of the sequence* $F_\omega, F_{\omega-1}, \ldots, F_{-\delta+1}$.

□

Many other possible definitions for matrix Padé approximants can be given. They will be published in another paper.

3. Solution of the scalar problem MPA(δ)

Suppose we want to solve problem MPA(δ) in the scalar case for $\delta = -\omega, \ldots, \omega$ with the normalizations introduced before. The trivial solution for $\delta = -\omega$ is given by

$$M_0 = 1\,, \quad N_0 = \sum_{k=0}^{\omega} F_k z^k\,.$$

We shall also introduce the residual

$$Y_0(z) = F(z)M_0(z) - N_0(z).$$

Since the F_k , $k > \omega$ do not appear in problem MPA(δ), we shall treat them as being zero, which implies that $Y_0(z) = 0$. This is not an essential assumption, but it will simplify the treatment a little bit.

We introduce an artificial $M_{-1} = L_0$ and an $N_{-1} = P_0$ with

$$P_0 = -z^{\omega+1} \quad \text{and} \quad L_0 = 0$$

and with corresponding residual

$$Y_{-1} = X_0 = F(z)L_0(z) - P_0(z) = z^{\omega+1}.$$

All these initializations are put into a global matrix

$$G_0(z) = \begin{bmatrix} X_0(z) & Y_0(z) \\ L_0(z) & M_0(z) \\ P_0(z) & N_0(z) \end{bmatrix}$$

for which we have the obvious relation

$$[-1 \;\; F(z) \;\; -1]G_0 = 0 \quad \text{with} \quad F(z) = \sum_{k=0}^{\omega} F_k z^k. \tag{3.1}$$

Because both X_0 and Y_0 are $O_+(z^{\omega+1})$, it follows that (1.7) of problem MPA(δ) is satisfied for both columns of G_0. If we multiply G_0 from the right with a polynomial matrix $V(z)$ to get $G(z) = G_0(z)V(z)$, then relation (3.1) will still hold for $G(z)$, and because the first row of $G(z)$ is given by $z^{\omega+1}[1 \;\; 0]V(z)$ we shall still satisfy condition (1.7) with the entries of $G(z)$. We shall use this fact to generate (unimodular) polynomial matrices V_n, $n = 1,2,\ldots$. They will be used to meet condition (1.8) while we do not have to pay attention to (1.7). So we define

$$G_{0,n}(z) = G_0(z)V_1(z)\cdots V_n(z) = G_0(z)V_{1,n}(z)$$
$$= \begin{bmatrix} X_{0,n}(z) & Y_{0,n}(z) \\ L_{0,n}(z) & M_{0,n}(z) \\ P_{0,n}(z) & N_{0,n}(z) \end{bmatrix}$$

where

$$\begin{bmatrix} X_{0,n}(z) & Y_{0,n}(z) \\ L_{0,n}(z) & M_{0,n}(z) \end{bmatrix} = \begin{bmatrix} z^{\omega+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1,n}(z) & Y_{1,n}(z) \\ L_{1,n}(z) & M_{1,n}(z) \end{bmatrix} = \begin{bmatrix} z^{\omega+1} & 0 \\ 0 & 1 \end{bmatrix} V_{1,n}(z).$$

Suppose that the minimal denominator degrees $\mu(\delta)$ satisfy

$$0 = \mu(-\omega) = \mu(-\omega+1) = \cdots = \mu(\delta_1-1) < \mu(\delta_1) = \cdots = \mu(\delta_2-1) < \mu(\delta_2) = \cdots,$$

i.e. the jumps in the denominator degree appear for the δ-values $\delta_1, \delta_2, \cdots$. The solutions of problem MPA(δ) will be the same between two successive δ_n. More precisely, the $[\omega,\delta]$ MPAs are the same for $\delta = \delta_n, \ldots, \delta_{n+1}-1$. The V_n matrices will be used to give a different solution at the jump points δ_n.

That δ_{n-1} and δ_n are two successive jump points means that

$$N_{0,n-1}M_{0,n-1}^{-1} = O_-(z^{-\delta})$$

for $\delta = \delta_{n-1}, \delta_{n-1}+1, \ldots, \delta_n-1$ where the order $-\delta$ is strict for the last δ-value δ_n-1. This means that

$$\partial N_{0,n-1} - \partial M_{0,n-1} = -(\delta_n-1). \tag{3.2}$$

This is also true for $n = 0$ if we set $\partial N_{-1} = \omega+1$ and $\partial M_{-1} = \mu(\delta_{-1}) = \mu(-\omega-1) = 0$.

Each V_n is made to satisfy $N_{0,n}(z)M_{0,n}(z)^{-1} = O_-(z^{-\delta_n})$ for increasing values of n but with the least possible degree $\mu(\delta_n)$ of $M_{0,n}(z)$.

The construction of a general V_n can be described as follows: suppose we have at a certain stage generated some $G_{0,n-1}$ with

1. $M_{0,n-1}(z)$ monic, of degree $\mu(\delta_{n-1})$. (3.3)

2. $M_{0,n-1}(z)^{-1}L_{0,n-1}(z) = O_-(z^{-1})$ (3.4)

3. $-P_{0,n-1}(z)$ is monic of degree $\omega - \mu(\delta_{n-1}) + 1$ (3.5)

4. $P_{0,n-1}(z)^{-1}N_{0,n-1}(z) = O_-(z^{-1})$. (3.6)

5. $N_{0,n-1}(z)M_{0,n-1}(z)^{-1} = O_-(z^{\omega-2\mu(\delta_{n-1})})$. (3.7)

Suppose the leading term in $N_{0,n-1}(z)$ is $Y_n z^{\eta_{n-1}}$. Then, because of (3.3)

$$N_{0,n-1}(z)M_{0,n-1}(z)^{-1} = Y_n z^{\eta_{n-1}-\mu(\delta_{n-1})} + \text{lower order terms}$$

with

$$\eta_{n-1} - \mu(\delta_{n-1}) = -\delta_n + 1.$$

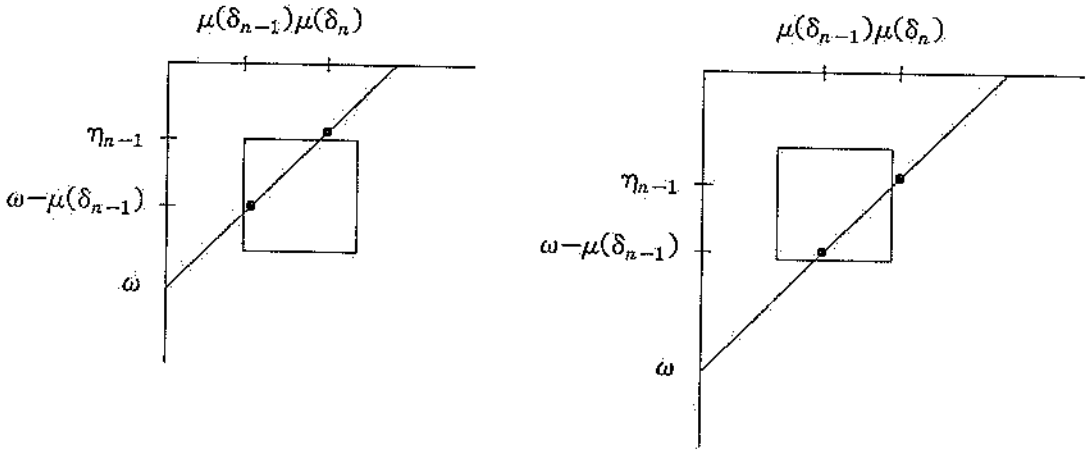
Now construct

$$G(z) = G_{0,n-1}(z)\begin{bmatrix} 0 & Y_n \\ L_n & z^{\beta} \end{bmatrix}$$

with

$$\beta = \partial P_{0,n-1} - \partial N_{0,n-1} = \omega - \mu(\delta_{n-1}) + 1 - \eta_{n-1} = \delta_n - 2\mu(\delta_{n-1}) > 0$$

and $L_n = -Y_n^{-1}$. It is well known in system theory [13,11] that this β will give the minimal degree $\mu(\delta_n)$ of $M(z)$ at the next jump point δ_n. This can be easily seen on the next picture.



Thus

$$\mu(\delta_n) = \mu(\delta_{n-1}) + \beta = \omega + \delta_n - \mu(\delta_{n-1}).$$

L_n will make $-P(z)$ monic. $M(z)$ is monic and $M(z)^{-1}L(z) = O_-(z^{-1})$ by construction. So (3.3-5) are satisfied for $G(z)$. We next try to satisfy (3.6). Therefore construct

$$G_{0,n}(z) = G(z)\begin{bmatrix} 1 & B_n(z) \\ 0 & 1 \end{bmatrix}$$

with $B_n(z)$ a polynomial of degree less than β such that

$$N_{0,n}(z) = P(z)B_n(z) + N(z) = O_-(z^{\eta_{n-1}-1}).$$

This will change $M(z)$ to make it satisfy (3.6) but leave its degree untouched while

$N_{0,n}(z)M_{0,n}(z)^{-1} = O_-(z^\gamma)$ with

$$\gamma = \eta_{n-1} - 1 - \mu(\delta_n) = \omega - 2\mu(\delta_n). \tag{3.8}$$

Thus (3.7) will be satisfied automatically. Thus the matrix $V_n(z)$ is given by

$$V_n(z) = \begin{bmatrix} 0 & Y_n \\ L_n & z^\beta \end{bmatrix} \begin{bmatrix} 1 & B_n(z) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & Y_n \\ L_n & z^\beta + L_n B_n(z) \end{bmatrix} = \begin{bmatrix} X_n & Y_n \\ L_n & M_n \end{bmatrix}.$$

A closer look at the calculations reveals that $-M_n(z)$ is in fact the quotient of $P_{0,n-1}(z)Y_n$ and $N_{0,n-1}(z)$, where the constant Y_n is chosen to make this quotient monic. Indeed, it follows from

$$P_{0,n-1}Y_n + N_{0,n-1}M_n = N_{0,n}$$

and

$$N_{0,n-1}L_n = P_{0,n}$$

that

$$\begin{aligned} N_{0,n-1}^{-1}[P_{0,n-1}Y_n] &= -M_n + N_{0,n-1}^{-1}N_{0,n} \\ &= -M_n + L_n P_{0,n}^{-1} N_{0,n} \end{aligned}$$

with the last term strictly proper (see (3.6)). Thus $-M_n$ is the quotient and $L_n N_{0,n}$ is the rest. Note also that (3.8) implies

$$\eta_n \leqslant \gamma + \mu(\delta_n) = \omega - \mu(\delta_n) \tag{3.9}$$

Example:

Take the example $F(z) = 1 + z^4 + z^5 + z^9 + \cdots$. Choose $\omega = 6$, the initial matrix G_0 is

$$G_0 = \begin{bmatrix} z^7 & 0 \\ 0 & 1 \\ -z^7 & 1+z^4+z^5 \end{bmatrix}, \quad N_{0,0}M_{0,0}^{-1} = \frac{1+z^4+z^5}{1}.$$

$$V_1 = \begin{bmatrix} 0 & 1 \\ -1 & z^2 \end{bmatrix} \begin{bmatrix} 1 & z-1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & z^2-z+1 \end{bmatrix}$$

Note that $Y_1 = -L_1 = 1$ and that M_1 is the quotient of z^7 divided by z^5+z^4+1. This gives

$$G_{0,1} = \begin{bmatrix} 0 & z^7 \\ -1 & z^2-z+1 \\ -z^5-z^4-1 & z^4+z^2-z+1 \end{bmatrix}, \quad N_{0,1}M_{0,1}^{-1} = \frac{1-z+z^2+z^4}{1-z+z^2}.$$

The next V-matrix is given by

$$V_2 = \begin{bmatrix} 0 & 1 \\ -1 & z \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1+z \end{bmatrix}$$

where $Y_2 = -L_2 = 1$ again and M_2 is the quotient of z^5+z^4+1 and z^4+z^2-z+1. So we find that

$$G_{0,2} = \begin{bmatrix} -z^7 & z^8+z^7 \\ -z^2+z-1 & z^3 \\ -z^4-z^2+z-1 & z^3 \end{bmatrix}, \quad N_{0,2}M_{0,2}^{-1} = \frac{z^3}{z^3}.$$

Next we find

$$V_3 = \begin{bmatrix} 0 & 1 \\ -1 & z \end{bmatrix}$$

and

$$G_{0,3} = \begin{bmatrix} -z^8-z^7 & z^9+z^8-z^7 \\ -z^3 & z^4-z^2+z-1 \\ -z^3 & -z^2+z-1 \end{bmatrix}, \quad N_{0,3}M_{0,3}^{-1} = \frac{-1+z-z^2}{-1+z-z^2+z^4}.$$

A next step gives

$$V_4 = \begin{bmatrix} 0 & -1 \\ 1 & z \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 1+z \end{bmatrix}$$

and

$$G_{0,4} = \begin{bmatrix} z^9+z^8-z^7 & z^{10}+2z^9+z^8 \\ z^4-z^2+z-1 & z^5+z^4-1 \\ -z^2+z-1 & -1 \end{bmatrix}, \quad N_{0,4}M_{0,4}^{-1} = \frac{-1}{-1+z^4+z^5}.$$

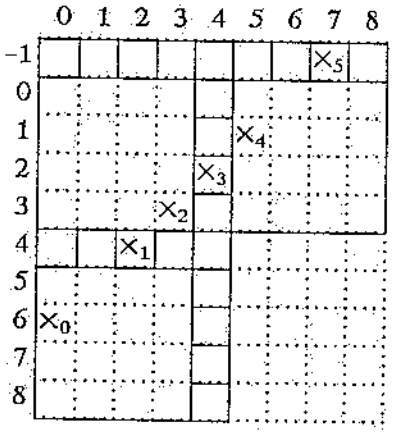
The last step gives

$$V_5 = \begin{bmatrix} 0 & 1 \\ 1 & -z^2 \end{bmatrix}\begin{bmatrix} 1 & 1-z \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & z^2-z+1 \end{bmatrix}$$

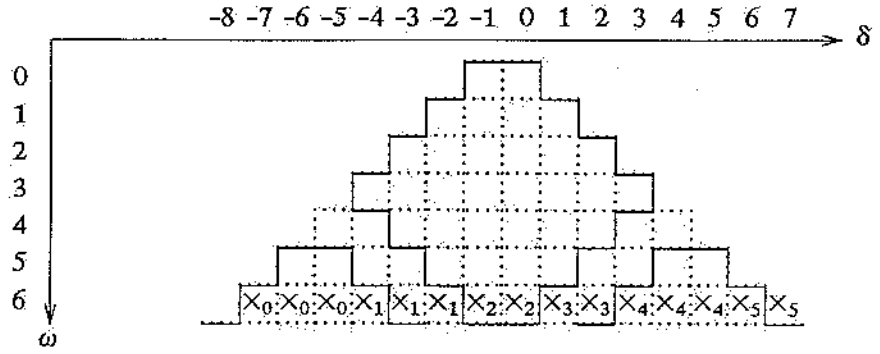
with

$$G_{0,5} = \begin{bmatrix} z^{10}+2z^9+z^8 & z^{12}+z^{11}+z^7 \\ z^5+z^4-1 & z^7 \\ -1 & 0 \end{bmatrix}$$

The successive solutions we obtained are indicated by an × in the Padé table of $F(z)$.



For the $[\omega,\delta]$ table this looks like



4. Solution of the matrix problem MPA(δ)

As for the scalar case, we suppose for simplicity that $F_k = 0$ for $k > \omega$ and we shall start with the initialization

$$G_0(z) = \begin{bmatrix} X_0(z) & Y_0(z) \\ L_0(z) & M_0(z) \\ P_0(z) & N_0(z) \end{bmatrix} = \begin{bmatrix} z^{\omega+1}I_p & 0 \\ 0 & I_m \\ -z^{\omega+1}I_p & N_0(z) \end{bmatrix}$$

with

$$N_0(z) = \sum_{k=0}^{\omega} F_k z^k$$

and define

$$G_{0,n}(z) = G_0(z)V_1(z)\cdots V_n(z) = G_0(z)V_{1,n}(z)$$
$$= \begin{bmatrix} X_{0,n}(z) & Y_{0,n}(z) \\ L_{0,n}(z) & M_{0,n}(z) \\ P_{0,n}(z) & N_{0,n}(z) \end{bmatrix}$$

where

$$\begin{bmatrix} X_{0,n}(z) & Y_{0,n}(z) \\ L_{0,n}(z) & M_{0,n}(z) \end{bmatrix} = \begin{bmatrix} z^{\omega+1}I_p & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} X_{1,n}(z) & Y_{1,n}(z) \\ L_{1,n}(z) & M_{1,n}(z) \end{bmatrix} = \begin{bmatrix} z^{\omega+1}I_p & 0 \\ 0 & I_m \end{bmatrix} V_{1,n}(z).$$

We have as before the relation

$$[-I_p \;\; F(z) \;\; -I_p]G_{0,n}(z) = 0$$

for all $n = 0,1,\ldots$ so that again condition (2.1) is satisfied for all $N_{0,n}$ and $M_{0,n}$ and therefore (2.1) need not be of our concern in what follows. Suppose now that the jumps in the minimal degree of the determinant of the denominator occur for $\delta = \delta_1, \delta_2, \ldots$. I.e.

$$0 = |\underline{\mu}(-\omega)| = \cdots = |\underline{\mu}(\delta_1 - 1)| < |\underline{\mu}(\delta_1)| = \cdots = |\underline{\mu}(\delta_2 - 1)| < |\underline{\mu}(\delta_2)| = \cdots .$$

Note that $|\underline{\mu}(\delta)| = |\underline{\lambda}(\delta)|$, so that the jumps for the $|\underline{\lambda}(\delta)|$ will occur at exactly the same places. The solution of problem MPA(δ) will not change between two successive δ_n.
To describe the construction of the unimodular V_n matrices that have to generate this change,

we start from the assumptions, analog to (3.3-7), which $G_{0,n-1}$ should satisfy :

1. $M_{0,n-1}(z)$ is CURI canonical, with column (and row) degrees $\underline{\mu}(\delta_{n-1})$. (4.1)
2. $M_{0,n-1}(z)^{-1}L_{0,n-1}(z) = O_{-}(z^{-1})$ (4.2)
3. $-P_{0,n-1}(z)$ is CLRI canonical with column (and row) degrees $\omega - \underline{\lambda}(\delta_{n-1}) + 1$ (4.3)
4. $P_{0,n-1}(z)^{-1}N_{0,n-1}(z) = O_{-}(z^{-1})$. (4.4)
5. $N_{0,n-1}(z)M_{0,n-1}(z)^{-1} = O_{-}(z^{\omega-\mu_j(\delta_{n-1})-\lambda_i(\delta_{n-1})})$. (4.5)

The notion of CLRI canonical is almost the same as CURI canonical, except that in the definition of CURI canonical we replace unit Upper triangular by unit Lower triangular. More precisely

Definition (CLRI canonical)

We call a polynomial matrix $P(z) \in \mathbb{C}^{p\times p}[z]$ CLRI canonical with degrees η_j, $j = 1, \ldots, p$ if it satisfies

1. The j-th column of $P(z)$ has degree η_j, $j = 1, \ldots, p$. (4.6)
2. The i-th row of $P(z)$ has degree η_i, $i = 1, \ldots, p$. (4.7)
3. The leading column coefficient matrix P_{hc} is unit lower triangular. (4.8)
4. The leading row coefficient matrix P_{hr} is the unit matrix. (4.9)

That these row and column degrees are related to the dual Kronecker indices as in (4.3) is because $P_{0,n}(z)$ is related to a left matrix fraction description. This will be explained in section 9.

Now the following theorem can be stated.

Theorem 4 : *Suppose the matrix $G_{0,n-1}(z)$ is associated with the unique solution of problem MPA(δ) for a given value of ω at the jump point δ_{n-1} and that it satisfies the conditions (4.1-5). Then there exists a unimodular matrix V_n^1 such that in*

$$G(z) = G_{0,n-1}(z)V_n^1 = \begin{bmatrix} X(z) & Y(z) \\ L(z) & M(z) \\ P(z) & N(z) \end{bmatrix}$$

the polynomials $M(z)$ and $N(z)$ satisfy conditions (2.1-3) of problem MPA(δ) at the next jump point $\delta = \delta_n$ and such that

1. *$M(z)$ is column reduced and* $\partial \det M(z) = |\underline{\mu}(\delta_n)|$ (4.10)
2. *$P(z)$ is column reduced with column degrees of $P(z)$ and $L(z)$ satisfying*
   $\partial P_j - \partial L_j > -\delta_n$. (4.11)

Proof : See appendix B. □

An algorithm describing the construction of V_n^1 is given in appendix E. It is called Generate_a_solution.

There may however be many solutions that can be generated in this way. By satisfying (4.1-5) for δ_{n-1} replaced by δ_n, we can make V_n unique.

Theorem 5 : *Let $G(z)$ be as in the previous theorem, then there exists a unimodular matrix V_n^2 such that*

$$G_{0,n} = G(z)V_n^2 = G_{0,n-1}(z)V_n^1V_n^2 = G_{0,n-1}(z)V_n$$

satisfies (4.1-5) for δ_{n-1} replaced by δ_n (i.e. (C.1-5) in appendix C). This $G_{0,n}(z)$ will be unique. Consequently also the transformation matrices V_n will be unique. It contains $N_{0,n}$ and $M_{0,n}$, which are the numerator and the unique CURI canonical denominator of the solution of problem MPA(δ) at the jump point $\delta = \delta_n$ as described in theorem 2. $N_{0,n}(z)M_{0,n}(z)^{-1}$ is the $[\omega,\delta]$ MPA for $F(z)$ for $\delta = \delta_n, \ldots, \delta_{n+1}-1$.

Proof : See appendix C.

□

The proofs of these theorems are constructive so that we have an algorithm to do the computations. The algorithm for the matrix case is considerably more involved than the one for the scalar case. A detailed description of the algorithm is given in the appendix E. When applied to a problem with $p = m = 1$, it reduces to the one given in the previous section. We shall apply it in the following example.

Example:

Consider the following simple example

$$F(z) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} z + \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} z^2 + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} z^3 + \cdots$$

and set $\omega = 3$. Then we get

$$\begin{bmatrix} X_0 & Y_0 \\ L_0 & M_0 \end{bmatrix} = \left[\begin{array}{cc|ccc} z^4 & 0 & 0 & 0 & 0 \\ 0 & z^4 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

and

$$[P_0 \;\; N_0] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{array}\right] z^4 + \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] z^3 + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] z^2 +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] z + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

By satisfying (2.2) and subsequent normalization we get

$$V_1 = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & z & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & z \end{array}\right] \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \hline -1 & 0 & z & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & z \end{array}\right]$$

which gives

$$\begin{bmatrix} X_{0,1} & Y_{0,1} \\ L_{0,1} & M_{0,1} \end{bmatrix} = \begin{bmatrix} z^4 I_2 & 0 \\ 0 & I_3 \end{bmatrix} V_1$$

and

$$[P_{0,1} \;\; N_{0,1}] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{array}\right] z^3 + \left[\begin{array}{cc|ccc} 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 \end{array}\right] z^2 +$$

$$\left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] z + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The last step gives

$$V_2 = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & z^2 & z \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ z & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ z & -1 & 0 & z^2 & z \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

which gives

$$\begin{bmatrix} X_{0,2} & Y_{0,2} \\ L_{0,2} & M_{0,2} \end{bmatrix} = \left[\begin{array}{cc|ccc} 0 & 0 & z^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & z^4 \\ \hline -1 & 0 & z & 0 & -1 \\ z & -1 & 0 & z^2 & z \\ 0 & 0 & -1 & -1 & z \end{array}\right]$$

and

$$[P_{0,2} \;\; N_{0,2}] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^3 + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^2 + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{array}\right] z + \left[\begin{array}{cc|ccc} 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

5. The scalar Euclidean algorithm

Given two polynomials $P(z)$ and $N(z)$, with $\partial P(z) \geqslant \partial N(z)$. Euclid's algorithm will compute recursively after the initialization

$$P_{0,0}(z) = P_0(z) = P(z), \qquad N_{0,0}(z) = N_0(z) = N(z)$$

the following divisions for $n = 1,2,\ldots$

$$P_{0,n-1}(z) = -M_n(z) N_{0,n-1}(z) + N_{0,n}(z), \quad \partial N_{0,n}(z) < \partial N_{0,n-1}(z)$$

$$P_{0,n}(z) = N_{0,n-1}(z)$$

which can also be written as

$$[P_{0,n}(z) \;\; N_{0,n}(z)] = [P_{0,n-1}(z) \;\; N_{0,n-1}(z)]\, V_n(z)$$

with

$$V_n(z) = \begin{bmatrix} 0 & 1 \\ 1 & M_n(z) \end{bmatrix}$$

and $-M_n(z)$ the quotient of $P_{0,n-1}(z)$ divided by $N_{0,n-1}(z)$.
The algorithm ends when $N_{0,n}(z) = 0$ and $P_{0,n}(z)$ will be a greatest common divisor (gcd) of $P_0(z)$ and $M_0(z)$.
If we want to obtain a unique gcd, we can choose the monic one. More generally, we can make every $M_{0,n}(z)$ monic by replacing $V_n(z)$ by a matrix of the form

$$V_n(z) = \begin{bmatrix} X_n & Y_n \\ L_n & M_n(z) \end{bmatrix}$$

with $X_n = 0$ and L_n an arbitrary nonzero constant. $-M_n(z)$ is the polynomial quotient of $P_{0,n-1}(z)Y_n$ and $N_{0,n-1}(z)$ with Y_n a nonzero constant which is chosen such that $M_n(z)$ is monic. The presence of L_n is an extra freedom which we shall use to make $-P_{0,n}(z) = -N_{0,n-1}(z)L_n$ monic.
If we define

$$V_{1,n} = \begin{bmatrix} X_{1,n}(z) & Y_{1,n}(z) \\ L_{1,n}(z) & M_{1,n}(z) \end{bmatrix} = V_1(z)V_2(z)\cdots V_n(z)$$

then

$$[P_{0,n}(z) \;\; N_{0,n}(z)] = [P_0(z) \;\; N_0(z)]V_{1,n}(z)$$

and if $N_{0,n}(z) = 0$, it follows that the monic gcd of $P_0(z)$ and $N_0(z)$ is given by

$$P_{0,n} = P_0(z)X_{1,n}(z) + N_0(z)L_{1,n}(z).$$

Note the following properties of the $V_n(z)$ matrices : If we start with some $P_0(z)$ and $N_0(z)$ and the degree of N_0 is strictly less than the degree of P_0, then we shall have at every stage that $P_{0,n}(z)^{-1}N_{0,n}(z) = O_-(z^{-1})$. Consequently $M_n(z)$ has a degree at least equal to 1 and since L_n is a constant, also $M_n(z)^{-1}L_n = O_-(z^{-1})$ (compare with (4.2)).

It is clear from what we have said before that the algorithm presented in section 3 to solve problem MPA(δ) is exactly the Euclidean algorithm applied to the polynomials $P_0(z) = -z^{\omega+1}$ and $N_0(z) = \sum_{k=0}^{\omega} F_k z^k$. Note that in general the gcd is obtained by adding one more step to the algorithm of section 3, i.e. for $\delta = \omega+1$.

In the example of section 3, the solution for $\delta = \omega$ was already a solution for $\delta = \omega+1$, so the normalized gcd of P_0 and N_0 is -1, since P_0 and N_0 are indeed coprime.

The ideas of relating the Euclidean algorithm and Padé approximation has been given e. g. in [14].

6. The matrix Euclidean algorithm

Since the algorithm of section 4 reduces in the scalar case to the algorithm of section 3, which we have identified as the Euclidean algorithm, it is natural to call the algorithm of section 4 a matrix Euclidean algorithm. In section 4 however it is only explained how the algorithm works for the special case that $P_0(z) = I_p z^{\omega+1}$. We shall have to define a proper extension to it for the case where $P_0(z) \neq I_p z^{\omega+1}$. But let us first recall the definition of a greatest common left divisor of two polynomial matrices.

Definition gcld : Suppose we have two polynomial matrices $P(z) \in \mathbb{C}^{p\times p}[z]$ and $N(z) \in \mathbb{C}^{p\times m}[z]$. Then a greatest common left divisor (gcld) is defined as a polynomial matrix $R(z) \in \mathbb{C}^{p\times p}[z]$ which is such that

1. $R(z)$ is a common left divisor. I.e. there exist polynomials $\overline{P}(z)$ and $\overline{N}(z)$ such that

$$P(z) = R(z)\overline{P}(z) \quad \text{and} \quad N(z) = R(z)\overline{N}(z). \tag{6.1}$$

2. It is a *greatest* common left divisor. I.e. if $S(z)$ is any other common left divisor, then there exists a polynomial $\overline{R}(z)$ such that

$$R(z) = S(z)\overline{R}(z). \tag{6.2}$$

The following theorem can be found in [12].

Theorem 6 : *Suppose we have two polynomial matrices $P(z) \in \mathbb{C}^{p\times p}[z]$ and $N(z) \in \mathbb{C}^{p\times m}[z]$. If there exists a unimodular matrix V such that*

$$[P(z)\ \ N(z)]\,V = [R(z)\ \ 0_{p\times m}]$$

then $R(z)$ will be a gcld of $P(z)$ and $N(z)$. Moreover, if $[P(z)\ \ N(z)]$ has full row rank, then all gcld will be nonsingular and they can only differ by unimodular (right) factors.

□

This theorem has the following corollary.

Corollary 1 : *Let the polynomial matrices $P(z)$ and $N(z)$ be as above and suppose that $[P(z)\ \ N(z)]$ has full row rank. Then there always exists a unimodular matrix V such that $[P_0(z)\ \ N_0(z)] = [P(z)\ \ N(z)]V$ satisfies the conditions (4.3-4). Consider the matrix*

$$G_0 = \begin{bmatrix} X_0 & Y_0 \\ L_0 & M_0 \\ P_0 & N_0 \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \\ P_0 & N_0 \end{bmatrix}.$$

When the algorithm of section 4 is applied to these data to generate the unimodular transformation matrices $V_1, \ldots, V_n$ then eventually, when $N_{0,n}(z) = 0$, $P_{0,n}(z)$ will be a gcld of $P(z)$ and $N(z)$ and it is the unique one whose negative is CLRI-canonical.

Proof (sketch) : Suppose that $[P(z)\ \ N(z)]$ has full row rank and set

$$G_0 = \begin{bmatrix} X_0 & Y_0 \\ L_0 & M_0 \\ P_0 & N_0 \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \\ P & N \end{bmatrix}.$$

For this matrix, (4.1) and (4.2) are clearly satisfied. If (4.3) and (4.4) are not satisfied, it is always possible to find a unimodular matrix V such that

$$[P'(z)\ \ N'(z)] = [P(z)\ \ N(z)]\,V$$

is as required, e.g., by algorithms like in appendix E. These P' and N' matrices are then used as initializations instead of P and N in the algorithm Initialization of appendix E. We can then apply the rest of the algorithm and we will find for $\delta = \omega+1$ (or for a smaller δ if it gives the same solution as for $\delta = \omega+1$) that in

$$G_{0,n} = G_0 V_{1,n} = \begin{bmatrix} X_{0,n} & Y_{0,n} \\ L_{0,n} & M_{0,n} \\ P_{0,n} & N_{0,n} \end{bmatrix}$$

$N_{0,n} = 0$. Since then

$$[P_0 \;\; N_0]\, V_{1,n} = [P_{0,n} \;\; 0]$$

with $V_{1,n}$ unimodular, we have found a gcld of P_0 and N_0. It is normalized (and therefore unique) as a CLRI canonical polynomial matrix with a sign change.

□

In our matrix example we found that the gcld of

$$P(z) = \begin{bmatrix} -z^4 & 0 \\ 0 & -z^4 \end{bmatrix} \quad \text{and} \quad N(z) = \begin{bmatrix} z^3+z & 1 & z^2 \\ z^2 & z & z^3 \end{bmatrix}$$

is given by

$$P_{0,2} = \begin{bmatrix} -z^3 & -1 \\ 0 & -z \end{bmatrix}$$

and we have indeed

$$P(z) = \begin{bmatrix} -z^3 & -1 \\ 0 & -z \end{bmatrix} \begin{bmatrix} z & -1 \\ 0 & z^3 \end{bmatrix} \quad \text{and} \quad N(z) = \begin{bmatrix} -z^3 & -1 \\ 0 & -z \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ z & -1 & -z^2 \end{bmatrix}.$$

The polynomial matrices $X_{0,n} = X_{1,n}$ and $L_{0,n} = L_{1,n}$ from $G_{0,n}$ appear in the relation expressing the gcld in terms of the given polynomials:

$$P_{0,n} = P_0 X_{1,n} + N_0 L_{1,n}.$$

7. Scalar continued fractions

It is well known that a continued fraction

$$\frac{A_0}{B_0}\Big] + \frac{A_1}{B_1}\Big] + \frac{A_2}{B_2}\Big] + \cdots$$

has convergents

$$\frac{A_{0,0}}{B_{0,0}} = \frac{A_0}{B_0}, \; \frac{A_{0,1}}{B_{0,1}}, \; \frac{A_{0,2}}{B_{0,2}}, \; \cdots$$

which satisfy the forward recurrence

$$A_{0,n} = A_{0,n-1} B_n + A_{0,n-2} A_n$$

$$B_{0,n} = B_{0,n-1} B_n + B_{0,n-2} A_n .$$

This recurrence is valid for $n = 1,2,\ldots$ if we set $A_{0,-1} = 0$ and $B_{0,-1} = 1$. It can also be expressed by

$$\begin{bmatrix} A_{0,n-1} & A_{0,n} \\ B_{0,n-1} & B_{0,n} \end{bmatrix} = \begin{bmatrix} A_{0,n-2} & A_{0,n-1} \\ B_{0,n-2} & B_{0,n-1} \end{bmatrix} \begin{bmatrix} 0 & A_n \\ 1 & B_n \end{bmatrix}$$

$$= \begin{bmatrix} 0 & A_0 \\ 1 & B_0 \end{bmatrix} \begin{bmatrix} 0 & A_1 \\ 1 & B_1 \end{bmatrix} \begin{bmatrix} 0 & A_2 \\ 1 & B_2 \end{bmatrix} \cdots \begin{bmatrix} 0 & A_n \\ 1 & B_n \end{bmatrix}.$$

A more general relation of the form

$$\begin{bmatrix} P_{0,n} & N_{0,n} \\ L_{0,n} & M_{0,n} \end{bmatrix} = \begin{bmatrix} P_0 & N_0 \\ L_0 & M_0 \end{bmatrix} V_{1,n} = \begin{bmatrix} P_0 & N_0 \\ L_0 & M_0 \end{bmatrix} V_1 V_2 \cdots V_n$$

with

$$V_i = \begin{bmatrix} X_i & Y_i \\ L_i & M_i \end{bmatrix}$$

represents a more general continued fraction of the form

$$\cfrac{N_0 + P_0 \cfrac{Y_1 + X_1 \cfrac{\cdots}{\cdots}}{M_1 + L_1 \cfrac{\cdots}{\cdots}}}{M_0 + L_0 \cfrac{Y_1 + X_1 \cfrac{\cdots}{\cdots}}{M_1 + L_1 \cfrac{\cdots}{\cdots}}}.$$

If we define the n-th convergent of this continued fraction to be the result we get when setting $P_n = L_n = 0$, then its convergents are given by $N_{0,n}M_{0,n}^{-1}$.

The Euclidean algorithm which solved problem MPA(δ) had

$$M_0 = 1, \quad N_0 = \sum_{k=0}^{\omega} F_k z^k, \quad P_0 = -z^{\omega+1}, \quad L_0 = 0 \text{ and } X_n = 0, \quad n = 1,2,\ldots,$$

and thus are the MPA-solutions $N_{0,n}M_{0,n}^{-1}$ the convergents of the continued fraction

$$N_0 - z^{\omega+1} \frac{Y_1}{M_1} + L_1 \frac{Y_2}{M_2} + L_2 \frac{Y_3}{M_3} + \cdots . \tag{7.1}$$

The denominator $M_{0,n}$ is monic and of degree $\mu(\delta_n)$ which is strictly increasing while the numerator degree ν is strictly decreasing and always bounded by $\omega - \lambda(\delta_n)$ with $\lambda(\delta_n) = \mu(\delta_n)$.

8. Matrix continued fraction

Once the more general continued fraction is associated with the general transformations V_i, it is not difficult to generalize everything to the matrix case.

We just mention that in the matrix version of section 4, the X_i matrices are not necessarily zero in general, so that instead of (7.1) we find that the $N_{0,n}M_{0,n}^{-1}$ are the convergents of the matrix continued fraction

$$N_0 - z^{\omega+1} I_p \cfrac{}{Y_1 + X_1 \cfrac{}{M_2 + L_2 \cfrac{}{\ddots}}} \cfrac{}{M_1 + L_1 \cfrac{}{Y_2 + X_2 \cfrac{}{M_2 + L_2 \cfrac{}{\ddots}}}}$$

where $\frac{A}{B}$ means AB^{-1}. Such a matrix continued fraction was first introduced by Antoulas [2].

9. Left/right duality

Everything that has been said can of course be translated in a left(right) dual result. We actually have used this duality in the proofs of the theorems. E.g. the Padé approximant has been defined as a right matrix fraction, but there also exists a left matrix fraction which approximates $F(z)$ in a Padé sense. These left MPAs can be obtained with dual algorithms, but the information needed is also present in the V_n matrices. Define

$$V_{1,n}^{-1} = W_{1,n} = \begin{bmatrix} T_{1,n} & -Q_{1,n} \\ -U_{1,n} & R_{1,n} \end{bmatrix}.$$

Define also

$$\Sigma_\omega = \sum_0^\omega F_k z^k,\ L_0' = 0_{m\times p}\ ,\ P_0' = I_p\ .$$

Then it follows from

$$\begin{bmatrix} L_{0,n}' & M_{0,n} \\ P_{0,n}' & N_{0,n} \end{bmatrix}^{-1} = \left\{ \begin{bmatrix} 0 & I_m \\ I_p & \Sigma_\omega \end{bmatrix} V_{1,n} \right\}^{-1} = W_{1,n} \begin{bmatrix} -\Sigma_\omega & I_p \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} -E_{0,n} & T_{1,n} \\ H_{0,n} & -U_{1,n} \end{bmatrix}$$

that

$$N_{0,n} M_{0,n}^{-1} = T_{1,n}^{-1} E_{0,n}\ .$$

I.e. the right matrix fraction can be written as a left matrix fraction with numerator

$$E_{0,n} = T_{1,n} \sum_{k=0}^{\omega} F_k z^k + Q_{1,n}$$

and denominator $D_{0,n} = T_{1,n}$. This left matrix fraction then is of course a left MPA. We can summarize this in the following theorem.

Theorem 7 : *With the notations introduced above, we find that a left MPA, solving the left MPA(δ_n) problem is given by $D_{0,n}^{-1} E_{0,n}$. The left denominator $D_{0,n}$ is the transpose of a CURI canonical matrix with row and column degrees $\lambda_i(\delta_n)$ if the $V_{1,n}$-matrix is associated with a right MPA solution that satisfies (4.1-5) for $\delta = \delta_n$.*

Proof : see appendix D. □

The converse can be shown along the same lines.

10. Conclusion

We have outlined how the scalar Euclidean algorithm can be generalized to the matrix case. It is normalized to give a unique continued fraction whose convergents are Padé approximants of $F(z)$ in the sense of the definition given in section 2 whenever we apply it to the data $-z^{\omega+1}I_p$ and $F(z) \in \mathbb{C}^{p\times m}[z]$, a polynomial of degree at most ω. Most important are the Kronecker and dual Kronecker indices of the leading coefficients in $F(z)$ since they describe the fine structure of the degrees of numerator and denominator of the approximants.

11. References

[1] P. ACHUTHAN, S. SUNDAR : A new application of the extended Euclidean algorithm for matrix Padé approximants. Comp. Math. Appl., 16, 287-296, 1988.

[2] B.M. ANDERSON, F.M. BRASCH Jr, P.V. LOPRESTI : The sequential construction of minimal partial realizations from finite input-output data. SIAM J. Control, 13, 552-571, 1975

[3] A.C. ANTOULAS: On recursiveness and related topics in linear systems, IEEE Trans. Automatic Control, AC-31(12),1121-1135, 1986.

[4] G.A. BAKER JR., P. GRAVES-MORRIS : *Padé approximants*, Addison-Wesley Publ. Comp., Reading, Massachusetts, 1981.

[5] Y. BISTRITZ: Nested bases of invariants for minimal realizations of finite matrix sequences, SIAM J. Control Opt., 21, 804-821, 1983.

[6] M.G. DE BRUIN: Continued fractions and a multi-dimensional Padé table. Thesis, Amsterdam University, 1974.

[7] B.W. DICKINSON, M. MORF, T. KAILATH : A minimal realization algorithm for matrix sequences. IEEE Trans. Autom. Control, AC-19, 31-38, 1974.

[8] A. DRAUX : Bibliography - Index - Report ANO-145, Université de Sciences et techniques de Lille, November 1984.

[9] P.A. FUHRMANN : A matrix Euclidean algorithm and matrix continued fraction expansion. Syst. Control Letters, 3, 263-271,1983.

[10] P. R. GRAVES-MORRIS : vector valued rational approximants II, I.M.A.J. Numer. Anal., 4, 209-224, 1984.

[11] W.B. GRAGG, A. LINDQUIST : On the partial realization algorithm, Lin. Alg. Appl., 50, 277-319, 1983.

[12] T. KAILATH: *Linear systems*, Prentice Hall, Englewood Cliffs, N.J., 1980.

[13] R.E. KALMAN : On partial realization, transfer functions and canonical forms, Acta Polytechnica Scandinavica, 31, 9-31, 1979.

[14] C.T. McELIECE, J.B. SHEARER : A property of Euclid's algorithm and an application to Padé approximation, SIAM J. Appl. Math., 34, 611-616, 1978.

[15] M. VAN BAREL, A. BULTHEEL : A Minimal partial algorithm for MIMO systems. Report TW79, July 1986, (Revised: April 1987) Department of Computer Science, K.U.Leuven.

[16] M. VAN BAREL, A. BULTHEEL : A Minimal partial algorithm for MIMO systems, II:A matrix continued fraction. Report TW91, June 1987, Department of Computer Science, K.U.Leuven.

[17] M. VAN BAREL, A. BULTHEEL : A Minimal partial algorithm for MIMO systems, III:A parametrization of all solutions. Report TW93, August 1987, Department of Computer Science, K.U.Leuven.

[18] M. VAN BAREL, A. BULTHEEL : A Minimal partial algorithm for MIMO systems, IV:A matrix Euclidean algorithm. Report TW94, December 1987, Department of Computer Science, K.U.Leuven.

[19] M. VAN BAREL, A. BULTHEEL : A Minimal partial algorithm for MIMO systems, V:A canonical matrix continued fraction. Report TW100, January 1988, Department of Computer Science, K.U.Leuven.

[20] G.L. XU : The existence and uniqueness of matrix Padé approximants. To appear Journ. Comput. Math.

[21] G.L. XU, J.K. LI : Generalized matrix Padé approximants. Submitted.

Appendix A : Proof of theorem 2

It will follow from the construction in section 4 that there always exists a solution with column degrees for the denominator equal to the Kronecker indices and with a matrix of leading column coefficients equal to a (unit) upper triangular matrix. Suppose $N(z)M(z)^{-1}$ is such a solution. Because it is a well established result that a polynomial matrix $M(z)$ has a unique CURI canonical form (see e.g. [12]), we will only prove that condition (2.4) will provide the uniqueness of the rational matrix function $N(z)M(z)^{-1}$.
It follows from the canonicity of $M(z)$ and (2.4) that

$$N(z) = O_-(z^{\omega-\lambda_i(\delta)-\mu_j(\delta)})M(z) = O_-(z^{\omega-\lambda_i(\delta)}). \tag{A.1}$$

If a right solution exists, then we can say that by duality also a left solution will exist (see section 9). Suppose it is $D(z)^{-1}E(z)$. For this right expression we have by duality that

$$E(z) = D(z)O_-(z^{\omega-\lambda_i(\delta)-\mu_j(\delta)}) = O_-(z^{\omega-\mu_j(\delta)}). \tag{A.2}$$

It follows from (A.1) and (A.2) that $EM = O_-(z^\omega)$ and also that $DN = O_-(z^\omega)$. Thus it follows that $EM - DN$ is a polynomial of degree at most ω. Now we also have that

$$F(z)M(z) - N(z) = O_+(z^{\omega+1}) \tag{A.3}$$

and

$$D(z)F(z) - E(z) = O_+(z^{\omega+1}). \tag{A.4}$$

Multiply (A.3) from the left with D and (A.4) from the right with M, then subtract and we find that

$$EM - DN = O_+(z^{\omega+1}).$$

Thus this expression must be zero. Therefore $NM^{-1} = D^{-1}E$. This can be shown for every right solution of MPA(δ). Since they will all be equal to the left $D^{-1}E$, they will be equal to each other. By duality also all the left solutions will be equal to each other and equal to the right solution.

□

Appendix B : Proof of theorem 4

Before we start the description of how to produce V_n^1, we shall give a basis for the space of all possible columns of M-vectors satisfying (2.1-3).
Denote by $H(\delta,\mu)$ the block Hankel matrix

$$H(\delta,\mu) = \begin{bmatrix} F_\omega & \cdots & F_{\omega-\mu} \\ \vdots & & \vdots \\ F_{\mu-\delta+1} & \cdots & F_{-\delta+1} \end{bmatrix}$$

and let $S(\delta,\mu)$ be its null space.

$$S(\delta,\mu) = \{m \in \mathbb{C}^{(\mu+1)p\times 1} | H(\delta,\mu)m = 0\}$$

For simplicity of notation, we shall make no distinction between $m(z) \in \mathbb{C}^{p\times 1}[z]$, a vector polynomial of degree μ and the vector m which contains its coefficients $m \in \mathbb{C}^{(\mu+1)p\times 1}$. Since the columns of $M(z)$ are selected from these null spaces with linear independent leading coefficients and minimal degrees, it is interesting to have a basis for all these null spaces. First we note the following simple property:

Lemma B.1 *With the notation introduced above, we have the following inclusion properties for the nullspaces :*

$S(\delta,\mu) \subseteq S(\delta+1,\mu+1)$
$S(\delta,\mu) \subseteq S(\delta,\mu+1)$
$zS(\delta,\mu) \subseteq S(\delta,\mu+1)$
$S(\delta,\mu) \subseteq S(\delta-1,\mu)$

Proof : This follows easily from the block Hankel structure of the matrix.

□

We shall denote the set of columns of degree μ for a particular choice of $M(z)$ satisfying (2.1-3) for a given δ by $\{\delta,\mu\}$. We can now describe a basis for the nullspace $S(\delta,\mu)$ in terms of these sets.

Lemma B.2 : *A basis for the null space $S(\delta,\mu)$ is given by*

$$B(\delta,\mu) = \bigcup_{i=0}^{\mu} B'(\delta-i,\mu-i)$$

with

$$B'(l,j) = \bigcup_{i=0}^{j} z^i\{l,j-i\}$$

Proof : The elements in

$$B'(\delta,\mu) = \{\delta,\mu\} \cup z\{\delta,\mu-1\} \cup \cdots \cup z^\mu\{\delta,0\}$$

have linearly independent leading coefficients and are all in $S(\delta,\mu)$. Hence, they can be chosen as part of $B(\delta,\mu)$. Any other basis vector in $B(\delta,\mu)\backslash B'(\delta,\mu)$ must have a coefficient of degree μ that is linearly dependent on the leading coefficients in $B'(\delta,\mu)$, otherwise it would be in a set $\{\delta,\mu-i\}$. Hence, it is always possible to replace a vector in $B(\delta,\mu)\backslash B'(\delta,\mu)$ by a vector of degree at most $\mu-1$ by subtracting a linear combination of $B'(\delta,\mu)$ vectors. Thus, by lemma A.2, it must be in $S(\delta-1,\mu-1)$. In other words :

$$B(\delta,\mu) = B'(\delta,\mu) \cup B(\delta-1,\mu-1)$$

The lemma then follows by induction.

□

Corollary B.1 : *A certain column of M can always be chosen so that its leading coefficient is a linear combination of the leading coefficients of the other columns of M that have a degree that is not higher.*

Proof : This follows immediately from the previous proof.

□

Now we turn to the proof of theorem 4. The proof is constructive and an algorithm to compute the matrix V_n^1 will be given implicitly. A detailed description is given in appendix E. To describe the V_n^1 matrix we shall introduce some elementary matrices representing elementary column operations for polynomial matrices.

$E_i(a)$ is the unit matrix with column i multiplied by a.
E_{ij} is the unit matrix with column i and column j interchanged
$E_{ij}(q(z))$ is the unit matrix with column j replaced by column j plus $q(z)$ times column i.

We shall describe how to change the original $G_{0,n-1}$-matrix of theorem 4 into the resulting G-matrix. So we shall start with a G-matrix initialized as $G_{0,n-1}$ (the initial stage) and describe the successive changes.

Furthermore, we introduce some notation to indicate certain column degrees. We denote the column degrees of $M(z)$ by $\underline{\mu} = \{\mu_1, \ldots, \mu_m\}$. Initially, these degrees are given by $\underline{\mu} = \underline{\mu}(\delta_{n-1}) = \underline{\mu}(\delta_n-1)$. Let us denote the column degrees of $P(z)$ by $\underline{\pi} = \{\pi_1, \ldots, \pi_p\}$ and the column degrees of $N(z)$ by $\underline{\eta} = \{\eta_1, \ldots, \eta_m\}$.

In general we shall denote in this proof the j-th column of an arbitrary matrix A by A_j.

A first problem is to detect the value of δ_n. Therefore we check for the successive values $\delta = \delta_{n-1}+1, \delta_{n-1}+2, \ldots$ if $\eta_i - \mu_i \leqslant -\delta$ is still satisfied for all the columns $i = 1,2, \ldots, p$. If so, we still have a solution for problem MPA(δ). The first time that this is not true anymore, we have found the value of δ_n. Suppose that for some i : $\eta_i - \mu_i > -\delta\ (= \delta_n)$. Since δ_n is the lowest possible δ-value for which this will happen, we have that at the previous value of δ

$$\eta_i - \mu_i = -\delta_n+1. \tag{B.1}$$

In that case something will have to be done. The idea is to eliminate the leading coefficient $HN_i = N_{\eta_i,i}$ of N_i. There are now two possiblities. This can be done by a unimodular transformation which either does not increase the denominator degree μ_i while η_i is decreased or we will have to increase the degree μ_i to a Kronecker index for $\delta = \delta_n$ but the numerator degree η_i is then increased with 1 less because the leading coefficient is eliminated. Thus in both cases $\eta_i - \mu_i \leqslant -\delta_n$ is satisfied again. We shall show that this can be obtained in the following way : Because the leading column coefficient matrix P_{hc} is nonsingular, we can write the leading coefficient as a linear combination of the columns of P_{hc}:

$$HN_i = -P_{hc}[c_1\ c_2\ \cdots\ c_p]^T .$$

Next choose an l such that $\pi_l = \max\{\pi_j | c_j \neq 0, j = 1,2, \ldots, p\}$. There are now two possibilities:

Case 1 : $\eta_i < \pi_l$. Then we set

$$V^{(i)} = E_{l,p+l}E_{p+i}(c_l)\prod_{j\neq l}E_{j,p+i}(c_j z^{\pi_l-\pi_j})E_{l,p+i}(z^{\pi_l-\eta_l}).$$

This will cause an increase in μ_i.
Case 2 : $\eta_i \geqslant \pi_l$. Then we set

$$V^{(i)} = \prod_{j=1}^{p}E_{j,p+i}(c_j z^{\eta_i-\pi_j}).$$

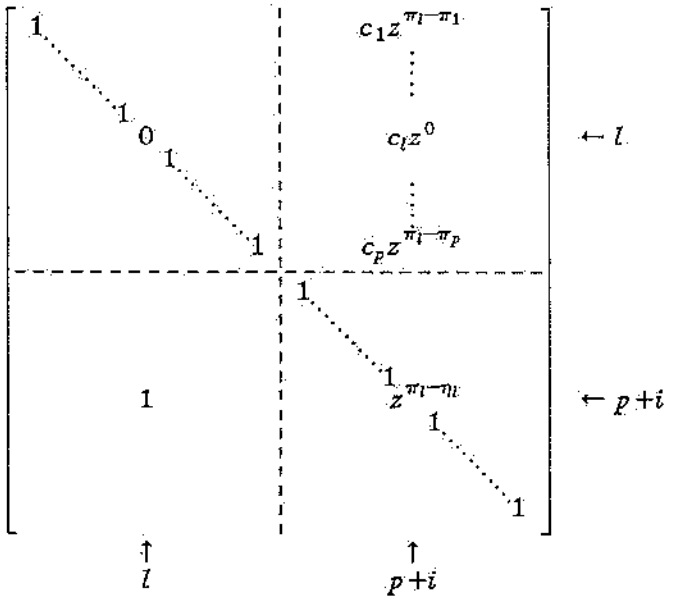
This will not increase μ_i. These two cases define $V^{(i)}$ matrices for all instances where $\eta_i - \mu_i > -\delta_n$. For the other columns where $\eta_i - \mu_i \leqslant -\delta_n$, nothing has to be done and we there set $V^{(i)} = I_{p+m}$. We now claim that

$$V_n^1 = \prod_{i=1}^{m}V^{(i)}$$

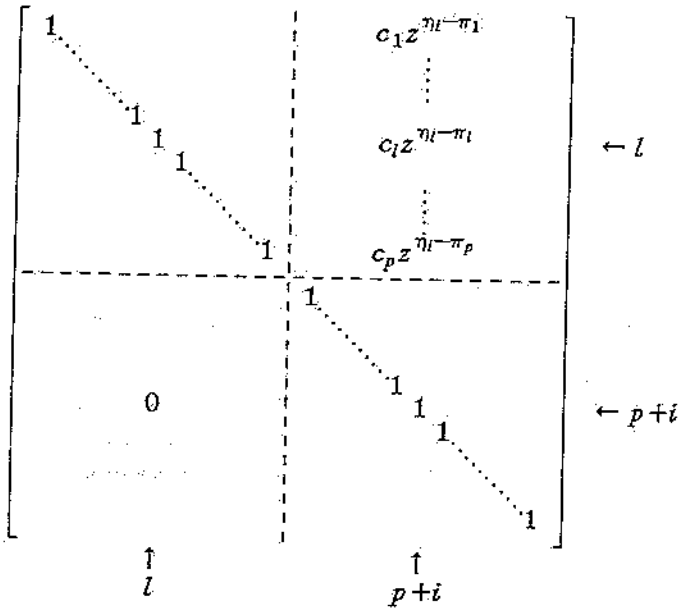
is the matrix needed in theorem 4.

1. The matrix V_n^1 is unimodular.

The matrix V_n^1 is certainly unimodular since each of the $V^{(i)}$ is unimodular and the product of unimodular matrices is unimodular. Each of the $V^{(i)}$ is unimodular since it has the form



in the first case and



in the second case. In both cases the determinant is equal to a constant.
Note that in the first case column l of the G-matrix will be replaced by column $p+i$ while there is no interchange in the second case.

2. *Condition (2.1) is satisfied (correct order of approximation).*

This is almost trivial to see. Note that

$$[-I_p \;\; F(z) \;\; -I_p]G_{0,n-1}V = 0 \tag{B.2}$$

whenever V is a unimodular matrix. Since (2.1) is satisfied for $G_{0,n-1}$, it follows that $[X_{0,n-1} \;\; Y_{0,n-1}] = O_+(z^{\omega+1})$ and since this is still true for $[X_{0,n-1} \;\; Y_{0,n-1}]V$ if V is polynomial, (2.1) remains satisfied.

3. *$P(z)$ is column reduced.*

This is another thing which is simple to verify. In case 2 $P(z)$ is not changed while in case 1 P_l is replaced by N_i but since the leading coefficient of N_i has by definition a nonzero component along the leading coefficient of P_l, it is clear that P_{hc} remains nonsingular, which means that $P(z)$ is column reduced.

4. *(4.11) i.e. $\partial P_j - \partial L_j > -\delta_n$ will be satisfied at every stage of the construction.*

To show this, we need the following lemma.

Lemma B.3 : *We have at the beginning of the construction of V_n^1 that*

$$\partial[P_{0,n-1}]_j - \partial[L_{0,n-1}]_j > -\delta_{n-1} \tag{B.3}$$

Proof : If $[L_{0,n-1}]_j = 0$ then $[P_{0,n-1}]_j$ will be the j-th column of the matrix $I_p z^{\omega+1}$ and the relation holds.
Suppose $[L_{0,n-1}]_j \neq 0$ and (B.3) is not true. Since $M_{0,n-1}$ is column reduced, its leading column coefficients form a regular matrix $M_{hc} = [M_{0,n-1}]_{hc}$ and it is thus possible to write the leading

coefficient of $[L_{0,n-1}]_j$ as a linear combination of M_{hc}. E.g. as

$$M_{hc}[c_1^T \ \cdots \ c_m^T]^T .$$

Since $[L_{0,n-1}]_j \neq 0$, there must be some k such that $c_k \neq 0$. For this k, it follows from (4.2) that $\partial[L_{0,n-1}]_j < \mu_k(\delta_{n-1})$. But in this case we may replace $[M_{0,n-1}]_k$ by $[L_{0,n-1}]_j$ because (2.2) is satisfied as we supposed and M_{hc} stays column reduced. Thus we still have a solution satisfying (2.1-3) but with

$$\partial[L_{0,n-1}]_j < \mu_k(\delta_{n-1})$$

which is in contradiction with the minimality of $M_{0,n-1}$ and the definition of the Kronecker indices.

□

Note that the columns of $P(z)$ and $L(z)$ are either replaced by columns from $N(z)$ and $M(z)$ or they are columns from the original $P_{0,n-1}(z)$ and $L_{0,n-1}(z)$. For the second possibility we have just proved that

$$\partial P_j - \partial L_j > -\delta_{n-1} > -\delta_n. \tag{B.4}$$

We even have for those columns

$$\partial P_j - \partial L_j > -\delta_n + 1. \tag{B.5}$$

If case 1 applies, then P_i and L_i are replaced by N_i and M_i, but it follows from (B.1) that for these also (4.11) holds.

5. Condition (2.2), i.e. $\partial N_i - \partial M_i < -\delta_n$ holds and $M(z)$ is column reduced.

The proof is by induction on i. If case 1 applies, then we had that (B.1) must hold before the transformation with $V^{(i)}$ is done. Also we have seen before that at every stage (4.11) holds. Thus

$$\partial L_k - \pi_k \leqslant \delta_n - 1. \tag{B.6}$$

Therefore we have in case 1 that the i-th column M_i of $M(z)$ is given by

$$z^{\pi_l - \eta_i} M_i + \sum_{k=1}^{p} c_k z^{\pi_l - \pi_k} L_k .$$

The first term has degree $\mu_i + \pi_l - \eta_i$ while the k-th term of the sum has degree $\pi_l - \pi_k + \partial L_k$. Now we have to consider two possibilities. Either P_k and L_k are still the original columns from $G_{0,n-1}$ or they have been replaced by some N_j and M_j column when case 1 occured at a smaller value of i. For the first possibility (B.5) holds and for the second possibility (B.6) holds. Thus

$$\pi_l - \pi_k + \partial L_k \leqslant \pi_l + \delta_n - 1 = \pi_l + \mu_i - \eta_i$$

with strict inequality if L_k has not been replaced. This means that the leading coefficient of the new i-th column M_i is equal to the original leading coefficient plus possibly some linear combination of leading coefficients of previous columns of the original M matrix. Since these were originally independent, they will stay so. Note also that in this case the degree of the new M_i will be

$$\mu_i^{new} = \mu_i(\delta_{n-1}) + \pi_l - \eta_i. \tag{B.7}$$

The degree of the new N_i will be

$$\eta_i^{new} \leqslant \eta_i + \pi_l - \eta_i - 1 = \pi_l - 1$$

since the leading coefficient is eliminated. So that these new degrees satisfy

$$\eta_i^{new} - \mu_i^{new} \leqslant \pi_l - 1 - \mu_i - \pi_l + \eta_i = \eta_i - \mu_i - 1 = -\delta_n$$

which means that (2.2) holds.
The second case is treated in a similar way. The only difference is that there is no increase in the degree of μ_i while the degree η_i is decreased by 1 because of the elimination.

6. The μ_i^{new} are the Kronecker indices $\mu_i(\delta_n)$.

The Kronecker indices are by definition the lowest possible obtainable degrees. Thus if we can prove that the degrees μ_i^{new} are the lowest possible ones, then the μ_i^{new} will be a permutation of the Kronecker indices and (2.3) will be satisfied. In the second place we shall show that they are the Kronecker indices, and not just a permutation thereof.

We shall distinguish between the L-vectors that are zero and those that are not. We shall call the nonzero L-vectors and the corresponding P-vectors the genuine L- and P-vectors respectively. The others will be called artificial. Note that an artificial P-vector has degree $\omega+1$ while all genuine P-vectors have degree ω or lower.

The minimality of the M-vectors is intertwined with a certain minimality property of the P-vectors. So we prove both minimality properties simultaneously.

Lemma B.4 : *Suppose $\delta = \delta_n$. The P-vectors are the vectors of lowest possible degree whose leading coefficients span the complete space $\mathbb{C}^p$ and that satisfy (2.1) and (2.2) for some value of $\delta < \delta_n$ but not for $\delta = \delta_n$.*

Proof : Suppose this is true at the initial stage i.e. for $\delta = \delta_n - 1$. We also have the minimality of the M-degrees for the value $\delta = \delta_n - 1$. We have to prove that the operations of the updating procedure do not disturb this property.
This is certainly true if the P-vectors of minimal degree are among the N-vectors, since we replace a P-vector by an N-vector (which corresponds to a value $\delta < \delta_n$) whenever it has a lower degree and can replace a P-vector without disturbing the column reducedness of the P-vectors. This is precisely the condition under which case 1 applies.
Suppose now that there exists a vector a of degree π which is not a column of $N_{0,n-1}$ but still satisfies (2.1-2) for a value $\delta < \delta_n$ and which can replace a P-vector while keeping P column reduced and yet lowering the degree of its determinant. Let b of degree β be the L-vector corresponding to a. Thus

$$\pi - \beta = -\sigma \geqslant -\delta_n + 1 \tag{B.8}$$

There are now two possibilities : either a replaces an artificial P-vector or a genuine P-vector.
An artificial vector can not be replaced by a because in that case it could have been replaced this a vector for $\delta = \sigma$ and this is in contradiction with the induction hypothesis which supposes the minimality of the P-vectors at all previous stages.
Suppose then that a replaces a genuine P-vector. This has to be a column of the $N_{0,n-1}$ matrix.

introduced at this stage, because if it were an N-column corresponding to a previous δ-value, then it could have been replaced at an earlier stage. Thus we suppose that a replaces the column of P which was the jth column of $N_{0,n-1}$. This had a degree η_j which satisfied

$$\eta_j - \mu_j(\delta_{n-1}) = -\delta_n + 1.$$

This together with $\pi < \eta_j$ and (B.8) gives

$$\beta = \pi + \sigma < \eta_j + \delta_n - 1 = \mu_j(\delta_{n-1}).$$

We distinguish the cases $\sigma \geqslant \delta_{n-1}$ and $\sigma < \delta_{n-1}$.
If $\sigma \geqslant \delta_{n-1}$, this is in contradiction with the minimality of the M-vectors because we could then replace $[M_{0,n-1}]_j$ and $[N_{0,n-1}]_j$ by b and a respectively and get a lower degree for $\det M_{0,n-1}$.
The case $\sigma < \delta_{n-1}$ can be reduced to this one. Therefore consider the stage where $\delta = \sigma+1 \leqslant \delta_{n-1} < \delta_n$. By induction hypothesis we know that the leading coefficient of a can be generated as a linear combination of the leading coefficients of the P-vectors for this δ-value. The degree π of a can not be smaller than the degrees of these P-vectors whose leading coefficients get a nonzero coefficient in this linear combination otherwise we could reduce the degree of the P-vectors and this contradicts their minimality at previous levels. Thus we can add polynomial linear combinations of those P-vectors to the a vector to eliminate its leading coefficient. Suppose this is obtained by the operation

$$a' = a + \sum_{k=1}^{p} c_k P_k z^{\pi - \partial P_k}.$$

The corresponding transformation of the b vector is then

$$b' = b + \sum_{k=1}^{p} c_k L_k z^{\pi - \partial P_k}.$$

Since by induction hypothesis the P- and L-degrees for the value $\delta = \sigma+1 < \delta_n$ satisfy

$$\partial P_k - \partial L_k \geqslant -\sigma,$$

we find that the degree of the terms in the sum which modifies b is bounded by $\pi - \partial P_k + \partial L_k \leqslant \pi + \sigma = \beta$. Thus the degree of b is not changed. For these new vectors a' and b' of degree $\pi' < \pi$ and $\beta' = \beta$ we have $\pi' - \beta' = -\sigma' \leqslant -\sigma - 1$. For this new σ we can repeat the same thing until $\sigma' \geqslant \delta_{n-1}$ and this reduces $\sigma < \delta_{n-1}$ to the case $\sigma \geqslant \delta_{n-1}$.

□

Now we come to the minimality of the M-vectors. Suppose there exists a vector u of degree $\mu < \mu_j^{new}$ that can replace the vector M_j. First note that $\mu \geqslant \mu_j(\delta_{n-1})$ and that M_j is not a column of $M_{0,n-1}$ because otherwise the matrix $M_{0,n-1}$ would not be minimal. Suppose v of degree η is the numerator vector corresponding to u. Because we have a solution for $\delta = \delta_n$, we know that

$$\eta - \mu \leqslant -\delta_n. \tag{B.9}$$

Now consider the vectors

$$u^* = u - z^{\mu - \mu_j(\delta_{n-1})}[M_{0,n-1}]_j \tag{B.10}$$

and

$$v^* = v - z^{\mu-\mu_j(\delta_n-1)}[N_{0,n-1}]_j. \tag{B.11}$$

You can easily verify that u^* has degree $\mu^* \leqslant \mu$ and that v^* has degree $\eta^* = \mu - \mu_j(\delta_{n-1}) + \eta_j > \eta$, where η_j is the degree of $[N_{0,n-1}]_j$. The leading coefficient of v is equal to the leading coefficient of $[N_{0,n-1}]_j$. We also know from corollary B.1 that the leading coefficient of u and hence also of u^* has only components along the leading coefficients of M-vectors with degree at most equal to μ. Thus it is possible to eliminate the leading coefficient of u^* by adding suitably shifted multiples of those M-columns. Of course we have to modify v^* correspondingly. The leading coefficient of v^* will not be affected by these operations. Suppose we call the result u^{**} and v^{**} with degrees $\mu^{**} = \mu^* - 1$ and $\eta^{**} = \eta^*$. These degrees now satisfy

$$\begin{aligned} \eta^{**} - \mu^{**} &\geqslant \eta^* - \mu^* + 1 \\ &\geqslant \mu - \mu_j(\delta_{n-1}) + \eta_j - \mu + 1 \\ &= -\delta_n + 2. \end{aligned} \tag{B.12}$$

This means that the denominator-numerator pair (u^{**},v^{**}) is a solution of (2.1-3) for a value of $\delta < \delta_n - 1$ and it is possible to eliminate the leading coefficient of $[N_{0,n-1}]_j$ with this vector. On the other hand, the algorithm needed a P-vector of degree $\pi_i = \eta_j^{new}+1$ to eliminate the same coefficient. Since we suppose by induction hypothesis that the P-vectors for $\delta = \delta_n - 1$ are minimal we find here a contradiction, so that $\mu \geqslant \mu_j^{new}$.

Now that we have proved that the μ_i^{new} are minimal and therefor a permutation of the Kronecker indices, we shall show that they are indeed the Kronecker indices, i.e. $\mu_i^{new} = \mu_i(\delta_n)$ for $i = 1, \ldots, m$.

This follows directly from the fact that M_{hc} is unit upper triangular. Indeed, if $[M_{hc}]_j$ has only nonzero entries for the rows $i \leqslant j$, then this means that the j-th column in the last block column of the block Hankel matrix of (2.5) can be written as a linear combination of the previous columns only. This, combined with the minimality, shows that the μ_j^{new} are indeed the Kronecker indices as defined in section 2.

□

An algorithm as it is described in the proof of this theorem is given in appendix E as Generate_a_solution.

Appendix C : Proof of theorem 5

First we note that the conditions (4.1-5) are redundant. We have the following Lemma :

Lemma C.1 : *Consider the conditions (4.1-5) for $\delta = \delta_n$, i.e. :*

1. $M_{0,n}(z)$ is CURI canonical, with column (and row) degrees $\underline{\mu}(\delta_n)$. (C.1)
2. $M_{0,n}(z)^{-1}L_{0,n}(z) = O_-(z^{-1})$ (C.2)
3. $-P_{0,n}(z)$ is CLRI canonical with column (and row) degrees $\omega - \underline{\lambda}(\delta_n) + 1$ (C.3)
4. $P_{0,n}(z)^{-1}N_{0,n}(z) = O_-(z^{-1})$. (C.4)
5. $N_{0,n}(z)M_{0,n}(z)^{-1} = O_-(z^{\omega-\mu_j(\delta_n)-\lambda_i(\delta_n)})$. (C.5)

Let $G(z)$ be a matrix as obtained by theorem 4. Suppose we can find a unimodular matrix V_n^2 such that $G_{0,n} = GV_n^2$ satisfying (C.1), (C.3) and (C.4), then it will also satisfy (C.5).

Proof : It follows from (C.4) and (C.3) that

$$\begin{aligned} N_{0,n}(z) &= P_{0,n}(z)O_-(z^{-1}) \\ &= O_-(z^{\omega-\lambda_i(\delta_n)+1})O_-(z^{-1}) \\ &= O_-(z^{\omega-\lambda_i(\delta_n)}). \end{aligned}$$

Using (C.1) we then find

$$\begin{aligned} N_{0,n}(z)M_{0,n}(z)^{-1} &= O_-(z^{\omega-\lambda_i(\delta_n)})O_-(z^{-\mu_j(\delta_n)}) \\ &= O_-(z^{\omega-\lambda_i(\delta_n)-\mu_j(\delta_n)}). \end{aligned}$$

Thus $M_{0,n}(z)$ and $N_{0,n}(z)$ also satisfy (C.5).

□

It will be shown in section 9 where the left/right duality is studied that if conditions (C.1-5) are satisfied for $G_{0,n}$, then we do not only fix a unique right solution for problem MPA(δ_n) with CURI canonical denominator with degrees $\underline{\mu}(\delta_n)$, but also a unique left solution with a denominator that is the transpose of a CLRI canonical (i.e. a CIRU canonical) polynomial matrix with degrees $\underline{\lambda}(\delta_n)$. This left solution is completely defined by $T_{1,n}(z)$ and $Q_{1,n}(z)$ where

$$W_{1,n} = V_{1,n}^{-1} = \begin{bmatrix} T_{1,n} & -Q_{1,n} \\ -U_{1,n} & R_{1,n} \end{bmatrix}.$$

Now suppose there are two matrices $G_{0,n} = G_{0,n-1}V_n$ and $G'_{0,n} = G_{0,n-1}V'_n$, both satisfying the conditions of theorem 5. We already know from theorem 2 that by conditions (C.1) and (C.5) we fix a unique representation of a unique right solution of MPA(δ). Thus this fixes the second block column of $G_{0,n}$. This implies $Y_{0,n} = Y'_{0,n}$, $M_{0,n} = M'_{0,n}$ and $N_{0,n} = N'_{0,n}$. Since now also the left solution is made unique and given a unique representation, we also have $T_{1,n} = T'_{1,n}$ and $Q_{1,n} = Q'_{1,n}$. Because

$$W'_{1,n}V_{1,n} = \begin{bmatrix} T_{1,n} & -U'_{1,n} \\ -U_{1,n} & R'_{1,n} \end{bmatrix} \begin{bmatrix} X_{1,n} & Y_{1,n} \\ L_{1,n} & M_{1,n} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ Z & I_m \end{bmatrix}$$

with

$$Z = R'_{1,n}L_{1,n} - U'_{1,n}X_{1,n}.$$

Hence

$$\begin{bmatrix} X_{1,n} & Y_{1,n} \\ L_{1,n} & M_{1,n} \end{bmatrix} = \begin{bmatrix} X'_{1,n} & Y_{1,n} \\ L'_{1,n} & M_{1,n} \end{bmatrix} \begin{bmatrix} I_p & 0 \\ Z & I_m \end{bmatrix} = \begin{bmatrix} X'_{1,n} + Y_{1,n}Z & Y_{1,n} \\ L'_{1,n} + M_{1,n}Z & M_{1,n} \end{bmatrix}.$$

Thus $L_{1,n} = L'_{1,n} + M_{1,n}Z$. But by (C.2) we get from this that $M^{-1}_{1,n}L_{1,n} = M^{-1}_{1,n}L'_{1,n} + Z$ and because $M_{1,n} = M_{0,n}$ and $L_{1,n} = L_{0,n}$ we finally get

$$M^{-1}_{0,n}L_{0,n} = M^{-1}_{0,n}L'_{0,n} + Z = O_-(z^{-1}).$$

Therefore $Z = 0$, which gives also $X_{1,n} = X'_{1,n}$ and $L_{1,n} = L'_{1,n}$ and consequently also $Y_{0,n} = Y'_{0,n}$ and $L_{0,n} = L'_{0,n}$ which implies $P_{0,n} = P'_{0,n}$. We thus have proved that $G_{0,n} = G'_{0,n}$.

□

Now that we have shown that $G_{0,n}$ will be unique whenever V^2_n exists, only this existence remains to be shown. The algorithm of appendix E describes the different steps to generate V^2_n. We shall not give the detailed proofs but give only a brief indication of what these algorithms do and why they work. We use 5 steps to realize conditions (C.1-4). Every step corresponds to a part of the algorithm given in appendix E.

1. Make_M_canonical

This algorithm makes M CURI canonical (C.1). It uses the fact that M_{hc} is already unit upper triangular. It makes $M_{hr} = I_m$ by eliminating the necessary coefficients without destroying the M_{hc}. It does not affect the column degrees, so minimality is maintained. The transformation is represented by a unimodular transformation on the right block column of G.

2. Make_P_{hc}_unit_lower_triangular

This algorithm realizes a part of the CLRI canonicity of $-P$. It makes P_{hc} unit lower triangular. The rest of this condition is realized in step 4. It takes a column which has a leading coefficient with a nonzero component on row i and whose degree is as low as possible. This column is brought to position i and the nonzero component at row i is normalized to -1. Then this component is eliminated in the leading coefficient of all the following columns. This does not change the column degrees of the P-columns and neither does it affect the span of their leading coefficients. Thus we do not destroy the property of the P-columns as described in lemma B.4. The transformation works only on the first block column of G and it is unimodular.

3. Make_$M^{-1}L$_strictly_proper

This algorithm realizes condition (C.2). It adds polynomial multiples of the second block column to the first block column. Therefore, it is a unimodular transformation. Also it does not affect the leading column coefficients of P because we have from (2.2)

$$\partial N_j \leq \partial M_j - \delta_n$$

and from (4.11)

$$\partial P_j > \partial L_j - \delta_n$$

4. Make_P_canonical

Now P is made CLRI canonical. This is done by an algorithm that is similar to Make_M_canonical. Since $-P_{hc}$ is now lower instead of upper unit triangular, the i-loop is not in reverse order as it was in Make_M_canonical. However by this operation, the condition realized in the previous step may be destroyed. To reinstall this, some extra work needs to be done. Note however that by this whole operation the property of lemma B.4 is not affected.

5. Make_P^{-1}N_strictly_proper

The final step realizes condition (C.4). This is similar to step 3. Again, as in the previous step the order of the i-loop is reversed because P is CLRI canonical while M was CURI canonical. Now a polynomial multiple of the first block column is added to the second one. This may destroy the achievement of step 1 and this has to be reinstalled by a unimodular operation on the second block column. It is easily verified that condition (2.2) is still satisfied.

This ends the proof of theorem 5.

□

Appendix D : Proof of theorem 7

Indeed, if (4.1,2,4,5) are satisfied with $\delta = \delta_n$, then $-P_{0,n}$ is CLRI canonical with degrees $\omega - \underline{\lambda}(\delta_n)+1$ iff $D_{0,n}^T$ is CURI canonical with column degrees $\underline{\lambda}(\delta_n)$. This follows from the fact that

$$P_{0,n} = \Sigma_\omega L_{1,n} - z^{\omega+1} X_{1,n}.$$

Since by (4.5)

$$\begin{aligned}\Sigma_\omega &= Y_{0,n} M_{0,n}^{-1} + O_-(z^{\omega-\lambda_i(\delta_n)-\mu_j(\delta_n)}) \\ &= z^{\omega+1} Y_{1,n} M_{1,n}^{-1} + O_-(z^{\omega-\lambda_i(\delta_n)-\mu_j(\delta_n)})\end{aligned}$$

and by (4.2)

$$L_{0,n} = L_{1,n} = O_-(z^{\mu_i(\delta_n)-1})$$

we find that

$$P_{0,n} = z^{\omega+1}(Y_{1,n} M_{1,n}^{-1} L_{1,n} - X_{1,n}) + O_-(z^{\omega-\lambda_i(\delta_n)-1}).$$

Furthermore, it follows from $V_{1,n} W_{1,n} = I_{p+m}$ that

$$T_{1,n}^{-1} = X_{1,n} - Y_{1,n} M_{1,n}^{-1} L_{1,n}.$$

So that

$$P_{0,n} = -z^{\omega+1} T_{1,n}^{-1} + O_-(z^{\omega-\lambda_i(\delta_n)-1}).$$

Because $-P_{0,n}$ is CLRI canonical with row degrees $\omega - \lambda_i(\delta_n)+1$ it follows that $T_{1,n}$ is column reduced with column degrees $\lambda_j(\delta_n)$ and with $(T_{1,n})_{hc}$ the unit matrix. On the other hand we have from (4.11) that the column degrees of $P_{0,n}$ and $L_{0,n}$ are related by

$$\partial(P_{0,n})_j > -\delta_n + \partial(L_{0,n})_j$$

and because

$$N_{0,n} M_{0,n}^{-1} = O_-(z^{-\delta_n}),$$

it follows from

$$P_{0,n} = -z^{\omega+1} T_{1,n}^{-1} + N_{0,n} M_{0,n}^{-1} L_{1,n}$$

that the leading column coefficient matrix of $P_{0,n}$ and of $-z^{\omega+1} T_{1,n}^{-1}$ are the same. Thus that $T_{1,n}$ is also row reduced with row degrees $\lambda_i(\delta_n)$ and with $(T_{1,n})_{hr}$ a unit lower triangular matrix. Hence $T_{1,n}^T$ is CURI canonical.

Appendix E : An algorithm

Canonical_MPA_solution :
{Compute the matrix G_n and V_n from G_{n-1} as in theorems 4 and 5
We use the following notations

$$G_{0,n} = G = \begin{bmatrix} X & Y \\ L & M \\ P & N \end{bmatrix}$$

with

$$L(z) = [L_{ij}] \text{ with } L_{ij} = \sum_{d=0}^{\alpha_{ij}} L_{i,j,d} z^d \ , \ L_{i,j,\alpha_{ij}} \neq 0$$

$$M(z) = [M_{ij}] \text{ with } M_{ij} = \sum_{d=0}^{\mu_{ij}} M_{i,j,d} z^d \ , \ M_{i,j,\mu_{ij}} \neq 0$$

$$P(z) = [P_{ij}] \text{ with } P_{ij} = \sum_{d=0}^{\pi_{ij}} P_{i,j,d} z^d \ , \ P_{i,j,\pi_{ij}} \neq 0$$

$$N(z) = [N_{ij}] \text{ with } N_{ij} = \sum_{d=0}^{\eta_{ij}} N_{i,j,d} z^d \ , \ N_{i,j,\eta_{ij}} \neq 0$$

Column degrees are defined by the following functions :

$$\mu_j = \max\{\mu_{ij} \mid i = 1, \ldots, m\}$$

$$\alpha_j = \max\{\alpha_{ij} \mid i = 1, \ldots, m\}$$

$$\pi_j = \max\{\pi_{ij} \mid i = 1, \ldots, p\}$$

$$\eta_j = \max\{\eta_{ij} \mid i = 1, \ldots, p\}$$

Furthermore we use the functions

$$\mu_{\min} = \min\{\mu_j \mid j = 1, \ldots, m\} \quad \text{and} \quad \pi_{\min} = \min\{\pi_j \mid j = 1, \ldots, p\}$$

We shall use the Gaussian-like elementary operations :

- $E_{i,j}$ = the unit matrix in which column i and column j are interchanged.
- $E_{i,j}(q)$ = the unit matrix in which column j is replaced by column j plus q times row i.
- $E_i(q)$ = the unit matrix where column i is multiplied by q.

We also use the convention that the degree of the zero polynomial (or Laurent series) is $-\infty$.}

Initialization

For $n = 0,1,...$

Generate_a_solution

Initialize_V_n^2

Make_M_canonical

Make_P_{hc}_unit_lower_triangular

Make_$M^{-1}L$_strictly_proper

Make_P_canonical

Make_$P^{-1}N$_strictly_proper

Define_V_n

end Canonical_MPA_solution

Initialization :

$$\text{Set } G = \begin{bmatrix} X & Y \\ L & M \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \\ -I_p & M(z) \end{bmatrix}$$

Set $\delta_{-1} = -1$

end Initialization

Generate_a_solution :
{Generate G and V_n^1 as in theorem 4}

Set $V_n^1 = I_{p+m}$

Set $\delta = \delta_{n-1}$

While $V_n^1 = I_{p+m}$

$\delta = \delta + 1$

For $i = 1, \ldots, m$

Set HN_i = the coefficient of $z^{\omega-\delta+\mu_i}$ in N_i

If $HN_i = 0$ then $V^{(i)} = I_{p+m}$

else $\{HN_i \neq 0\}$

Find the c_k such that $HN_i + P_{hc}[c_1, \ldots, c_p]^T = 0$

Choose some l such that $\pi_l = \max\{\pi_j | c_j \neq 0,\ j = 1, \ldots, p\}$

If $\pi_l > \eta_i$ then

$$V^{(i)} = E_{l,p+i}E_{p+i}(c_l)\prod_{\substack{j=1\\ j\neq l}}^{p} E_{j,p+i}(c_j z^{\pi_l-\pi_j})E_{l,p+i}(z^{\pi_l-\eta_i})$$

else $\{\pi_l \leqslant \eta_i\}$

$$V^{(i)} = \prod_{j=1}^{p} E_{j,p+i}(c_j z^{\eta_i-\pi_j})$$

Replace G by $GV^{(i)}$

Set $V_n^1 = \prod_{i=1}^{m} V^{(i)}$

end Generate_a_solution

Initialize_V_n^2 :

$V_n^2 = I_{p+m}$

end Initialize_V_n^2

Make_M_canonical :
{Make $M(z)$ CURI canonical (see C.1). Note that M_{hc} is already upper triangular}

For $j = 1,2,\ldots,m$

For $d = \mu_j, \mu_j - 1, \ldots, \mu_{\min}$

For $i = m, m-1, \ldots, 1$

If $i \neq j$ and $d \geqslant \mu_i$ and $M_{i,j,d} \neq 0$ then

$\gamma = d - \mu_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{p+i,p+j}(-M_{i,j,d} z^{\gamma})$

end Make_M_canonical

Make_-P_{hc}_unit_lower_triangular :
{Before satisfying condition (C.2) we make $-P_{hc}$ unit lower triangular.
Note that it is already column reduced}

For $i = 1, \ldots, p$

Choose some l such that $e = \pi_{il} = \min\{\pi_{ij} | \pi_{ij} = \pi_j,\ j = i, \ldots, p\}$

Set $q = -P_{i,l,e}^{-1}$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{i,l} E_i(q)$

For $j = i+1, \ldots, p$

If $\pi_i \leqslant \pi_j$ and $P_{i,j,\pi_j} \neq 0$ then

Set $\gamma = \pi_j - \pi_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{i,j}(P_{i,j,\pi_j} z^{\gamma})$

end Make_-P_{hc}_unit_lower_triangular

Make_M^{-1}L_strictly_proper:
{This realizes condition (C.2)}

For $j = 1,2,\ldots,p$

For $d = \alpha_j,\ldots,\mu_{\min}$

For $i = m,m-1,\ldots,1$

If $d \geqslant \mu_i$ and $L_{i,j,d} \neq 0$ then

$\gamma = d - \mu_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{p+i,j}(-L_{i,j,d} z^{\gamma})$

end Make_M^{-1}L_strictly_proper

Make_P_canonical :
{This realizes condition C.3 while keeping condition C.2}

For $j = 1, \ldots, p$

For $e = \pi_j, \pi_j - 1, \ldots, \pi_{\min}$

For $i = 1, \ldots, p$

If $i \neq j$ and $e \geqslant \pi_i$ and $P_{i,j,e} \neq 0$ then

$\gamma = e - \pi_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{i,j}(P_{i,j,e} z^{\gamma})$

{This part makes $M(z)^{-1}L(z)$ strictly proper again}

For $d = \alpha_j, \alpha_j - 1, \ldots, \mu_{\min}$

For $i = m, m-1, \ldots, 1$

If $d \geqslant \mu_i$ and $L_{i,j,d} \neq 0$ then

$\gamma = d - \mu_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{p+i,j}(-L_{i,j,d} z^{\gamma})$

end Make_P_canonical

Make_P^{-1}N_strictly_proper :
{This satisfies condition C.4 while C.1 is kept satisfied}

For $j = 1, \ldots, m$

For $e = \eta_j, \eta_j - 1, \ldots, \pi_{\min}$

For $i = 1, \ldots, p$

If $e \geqslant \pi_i$ and $P_{i,j,e} \neq 0$ **then**

$\gamma = e - \pi_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{i,p+j}(N_{i,j,e} z^{\gamma})$

{make $M(z)$ CURI canonical again}

For $d = \mu_j - 1, \mu_j - 2, \ldots, \mu_{\min}$

For $i = m, m-1, \ldots, 1$

If $d \geqslant \mu_i$ and $M_{i,j,d} \neq 0$ **then**

$\gamma = d - \mu_i$

Replace $\begin{bmatrix} G \\ V_n^2 \end{bmatrix}$ by $\begin{bmatrix} G \\ V_n^2 \end{bmatrix} E_{p+i,p+j}(-M_{i,j,d} z^{\gamma})$

end Make_P^{-1}N_strictly_proper

Define_V_n :

Set $V_n = V_n^1 V_n^2$

end Define_V_n