A matrix Euclidean algorithm and matrix Padé approximation*

A. Bultheel
M. Van Barel

Report TW 104, January 1988

Department of Computer Science, K.U.Leuven

ABSTRACT

We give a generalization of the Euclidean algorithm to the rectangular matrix case. The algorithm is normalized so that it generates unique matrix Padé approximants and a unique continued fraction representation for them.

* This paper was presented at the meeting on Methoden und Anwendungen der Approximationstheorie, Jan 31-Feb 6, 1988 in Oberwolfach.

A matrix Euclidean algorithm and matrix Padé approximation (†)

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

ABSTRACT

We give a generalization of the Euclidean algorithm to the rectangular matrix case. The algorithm is normalized so that it generates unique matrix 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 rational form $N(z)M(z)^{-1}$ can be simply normalized by requiring e.g. that $M(z)$ be monic. If however $N(z) \in \mathbb{C}^{p \times m}$ and $M(z) \in \mathbb{C}^{m \times m}$, the normalization is not simple. We have chosen for a canonical (in a certain sense) normalization for $M(z)$.

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, it turns out to be most useful to consider the degrees of the columns of the denominator and the row degrees of the numerator. In this way we get sufficiently fine information about the degrees of freedom in the rational form to be able to give a definition of a Padé approximant.

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. We shall give a matrix generalization of the Euclidean algorithm which will generate matrix Padé approximants which have the same property.

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

We shall not go into the details of the proofs, since this would be too long an exposition. Instead we shall give a survey by explaining what happens in the scalar case in a way that will make the matrix generalization plausible, although the proof may not be simple.

Matrix Padé approximants and more generally Padé approximants in a noncommutative algebra have been considered by many authors before us. Draux [7] has commented on about

(†) This paper was presented at the meeting on Methoden und Anwendungen der Approximationstheorie, Jan 31-Feb 6, 1988 in Oberwolfach.

300 references on this topic. The vector case i.e. ($m=1$) has been considered by de Bruin [5] and Graves Morris [9]. 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 [1] and [6]. In this field Fuhrmann [8] describes a matrix Euclidean algorithm and a matrix continued fraction. His results don't 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 [12-16] on algorithms for minimal partial realization in the matrix case, a paper of Antoulas [2] 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 [16] 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. In fact we shall give in this text a translation of the results of [16] in a Padé terminology, not using the system theoretic vocabulary of the partial realization literature.

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 $\{F_\omega, \ldots, F_{\omega-\delta}\}$
- $\underline{\lambda}(\delta) = \{\lambda_1(\delta), \ldots, \lambda_p(\delta)\}$: dual Kronecker indices of $\{F_\omega, \ldots, F_{\omega-\delta}\}$
- ∂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 Padé problem

The scalar Padé approximation problem is well known [3]. Given a formal power series $F(z) = \sum_{k=0}^{\infty} F_k z^k$, with $F_k \in \mathbb{C}$, we could define a $[\nu/\mu]$ Padé approminant (PA) as the rational form $N(z)M(z)^{-1}$ satisfying

1. $N(z)$ and $M(z)$ are coprime
2. $F(z) - N(z)M(z)^{-1} = O_+(z^{\omega+1})$
3. $N(z) = O_-(z^\nu)$
4. $M(z) = O_-(z^\mu)$
5. $\omega \geqslant \nu + \mu$.

The table containing at the intersection of row ν and column μ the $[\nu/\mu]$ 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 $[\nu/\mu]$ approximants we can modify the definition of a PA as follows:

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

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (1.1)
2. $N(z) = O_-(z^{\nu})$ (1.2)
3. $M(z) = O_-(z^{\mu})$ (1.3)
4. ∂M as low as possible (1.4)
5. $\omega \geqslant \nu + \mu$. (1.5)

The $[\nu/\mu]$ approximant can be normalized by imposing that $M(z)$ should be monic.
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	μ
0	1/1	1/1	1/1	1/1	
1	1/1	1/1	1/1		
2	1/1	1/1			
3	1/1				
ν					

First definition

	0	1	2	3	μ
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	
ν					

Definition PA

Note that both definitions are equivalent iff $M(0) \neq 0$.
If we are given the order ω, and ask which are the simplest $N(z)$ and $M(z)$ that satisfy (1.1) 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 $N(z)M(z)^{-1} = O_-(z^{\omega-\delta})$. So we consider the following problem for fixed $F(z)$ and ω :

Problem PA(δ) : Find polynomials $M(z)$ and $N(z)$ such that

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (1.6)
2. $N(z)M(z)^{-1} = O_-(z^{\omega-\delta}), \quad \delta = 0,1,\ldots,2\omega$ (1.7)
3. $M(z) = O_-(z^{\mu})$ with μ as low as possible. (1.8)

From (1.8) it follows that μ must be ∂M which should be as low as possible. If $M(z)$ has to satisfy (1.6) with $\partial M = \mu$ and $N(z) = O_-(z^{\nu})$, then the coefficients of $M(z)$ should satisfy the Hankel system

$$\sum_{l=0}^{\mu} F_{\omega-i-l} M_l = 0 \ , \quad i = 0,1,\ldots,\omega-\nu-1 \tag{1.9}$$

where $M(z) = \sum M_j z^j$.
To satisfy (1.7), we must have $N(z) = O_-(z^{\nu})$ with $\nu \leqslant \mu+\omega-\delta$. Thus to satisfy (1.6-8), we try first $\mu = 0$, then $\mu = 1$ etc. I.e. we successively try to solve the Hankel systems ($F_{-k} = 0$ for $k > 0$)

$$\begin{bmatrix} F_\omega & \cdots & F_{\omega-\mu} \\ \vdots & & \vdots \\ F_{\omega-\delta+\mu+1} & \cdots & F_{\omega-\delta+1} \end{bmatrix} \begin{bmatrix} M_0 \\ \vdots \\ M_\mu \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{1.10}$$

for $\mu = 0,1,\ldots,\delta$. For $\mu = \delta$ there is no system (1.10) so that $M(z)$ can be chosen to be an arbitrary polynomial of degree $\mu = \delta$ and also for $\delta = 0$, the system (1.10) doesn't 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 columns of the Hankel matrix of the systems (1.10) are linearly dependent will give the minimal μ for which a solution of problem PA(δ) exists. We represent this minimal μ as $\mu(\delta)$. As we shall see later, there will always exist a solution of problem PA(δ) with $\nu \leqslant \omega - \mu(\delta)$, so that $\nu + \mu(\delta) \leqslant \omega$, which means that this is a PA solution of problem PA(δ). If more solutions of problem PA(δ) exist, we can take the one with minimal numerator degree. This solution will be unique up to a constant multiple. Hence to fix the solution of PA(δ) completely, we can normalize it and require $M(z)$ to be monic.

A geometrical interpretation of the problem is as follows: Consider the set S_ω of all $[\nu/\mu]$ PA's with $\nu + \mu \geqslant \omega$. Every element in S_ω will satisfy condition (1.6). In the set S_ω drop all elements that are in a block whose left top element is below the diagonal $\nu - \mu = \delta - \omega$. The remaining elements will satisfy both conditions (1.6) and (1.7). Among these select those PA's for which μ is minimal and finally take the one with minimal ν. This element will be the one that solves problem PA(δ). If it is normalized properly, it is the unique solution described before.
Note that the elements above the main diagonal of a block will never be selected as a solution for problem PA(δ).
For our previous example, the entries of the block that solve problem PA(δ) for some values of ω and δ are given below.

	0	1	2	3	μ
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	
ν					

Solutions of problem PA(δ).

As you can see, not every PA will be a solution of problem PA(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.

2. The matrix Padé problem

We shall now give the natural generalization of the theory of the previous section for the case that the $F_k \in \mathbb{C}^{p\times m}$. In the extension of definition PA we now consider polynomial matrices $N(z) \in \mathbb{C}^{p\times m}[z]$ and $M(z) \in \mathbb{C}^{m\times m}[z]$. Also we shall consider a set of row degrees $\underline{\nu} = \{\nu_1,\ldots,\nu_p\}$ for $N(z)$ and a set of column degrees $\underline{\mu} = \{\mu_1,\ldots,\mu_m\}$ for $M(z)$. The definition is as follows :

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

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (2.1)
2. $N(z) = O_-(z^{\nu_i})$ (2.2)
3. $M(z) = O_-(z^{\mu_j})$ (2.3)
4. $\partial \det M$ as low as possible (2.4)

$$5.\quad p\omega \geqslant \sum_{i=1}^{p}\nu_i + \sum_{j=1}^{m}\mu_j\ . \tag{2.5}$$

Since the rows of $N(z)$ and the columns of $M(z)$ are m-vectors, we find that there are

$$m[\sum_{i=1}^{p}(\nu_i+1) + \sum_{j=1}^{m}(\mu_j+1)]$$

parameters in $N(z)$ and $M(z)$. Each F_k has pm entries. Condition (2.5) then is a way of saying that the number of conditions we find in (2.1) is at least equal to the number of free parameters left by (2.2) and (2.3).
Again we can look for the simplest rational form which satisfies the first of these conditions and impose a relative importance to numerator and denominator degree by requiring $N(z)M(z)^{-1} = O_-(z^{\omega-\delta})$. Therefore we consider problem PA(δ) which is almost as before with given $F(z)$ and ω, parametrized by δ.

Problem PA(δ): Find polynomials $M(z)$ and $N(z)$ such that

1. $F(z)M(z) - N(z) = O_+(z^{\omega+1})$ (2.6)
2. $N(z)M(z)^{-1} = O_-(z^{\omega-\delta}),\quad \delta = 0,1,\ldots,2\omega$ (2.7)
3. $M(z) = O_-(z^{\mu_j})$ with the μ_j as low as possible. (2.8)

As in the scalar case it follows that the column degrees of M are the μ_j which have to be as low as possible.
Let us denote the columns of $M(z)$ by

$$m_j(z) = \sum_{i=0}^{\mu_j} m_{ij} z^i.$$

Then condition (2.8) implicitly requires that $M(z)$ has a regular matrix of leading column coefficients, i.e.

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

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

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

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

$$\sum_{l=0}^{\mu_j} F_{\omega-i-l}m_{lj} = 0\ ,\quad i = 0,1,3,\ldots,\omega-\eta_j-1\ ,\quad j = 1,2,\ldots,m$$

where the η_j are defined by $N(z) = O_-(z^{\eta_j})$. Moreover we have to make M_{hc} nonsingular. Since condition (2.7) requires $\eta_j \leqslant \mu_j+\omega-\delta$, we can consider as in the scalar case the block Hankel systems (we suppose again that $F_{-k} = 0$ for $k > 0$)

$$\begin{bmatrix} F_\omega & \cdots & F_{\omega-\mu_j} \\ \vdots & & \vdots \\ F_{\omega-\delta+\mu_j+1} & \cdots & F_{\omega-\delta+1} \end{bmatrix} \begin{bmatrix} m_{0j} \\ \vdots \\ m_{\mu_j,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.9}$$

for $\mu_j = 0,1,\ldots,\delta$ as low as possible and M_{hc} regular. This problem is well known in the theory of block Hankel matrices and the minimal μ_j's are known as the Kronecker indices of the sequence $F_\omega, F_{\omega-1}, \ldots, F_{\omega-\delta+1}$. A practical scheme to construct these indices can be described as follows: For $j = 1,\ldots,m$ call h_j the j-th column of the matrix

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

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

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

and place h_{m+j} at entry $(1,j)$ of the column table. These columns have length $(\delta-1)p$. Continue like that, each time dropping the first block entry of the sequence until the m columns $h_{(\delta-1)m+j}$ of $F_{\omega-\delta+1}$ are placed at positions $(\delta-1,j)$ 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 $h_{im+j} \in \mathbb{C}^{(\delta-i)p}$ is linear independent of all preceding columns truncated at the same length, we place a $\times$ at entry (i,j), otherwise we place a $\bullet$.

	1	2	$\cdots$	m
0	h_1	h_2	$\cdots$	h_m
1	h_{m+1}	h_{m+2}	$\cdots$	h_{2m}
2	h_{2m+1}	h_{2m+2}	$\cdots$	h_{3m}
3	h_{3m+1}	h_{3m+2}	$\cdots$	h_{4m}
$\vdots$	$\vdots$	$\vdots$		$\vdots$

column table

	1	2	$\cdots$	m
0	$\times$	$\times$	$\cdots$	$\times$
1	$\bullet$	$\times$	$\cdots$	$\times$
2	$\bullet$	$\times$	$\cdots$	$\bullet$
3	$\bullet$	$\bullet$	$\cdots$	$\bullet$
$\vdots$	$\vdots$	$\vdots$		$\vdots$

crate table

Note 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 (2.9) has a solution. It is a property of this 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 also be dots. Now consider the j-th column of the crate table. If we call $\mu_j(\delta)$ the row index of the first dot appearing in that column, then these will be the minimal μ_j needed to solve problem PA(δ).
The dual Kronecker indices $\lambda_i(\delta)$, $i = 1,2, \ldots, p$ are defined in a similar way, now working with the rows of

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

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 PA(δ). However, up to a normalization, there is only one which satisfies [4]

$$N(z)M(z)^{-1} = O_-(z^{\omega-\lambda_i(\delta)-\mu_j(\delta)}). \tag{2.10}$$

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 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 $M(z)$ C1 canonical by which we mean that

1. The j-th column of $M(z)$ has degree $\mu_j(\delta)$, $j = 1, \ldots, m$, (2.11)
2. The i-th row of $M(z)$ has degree $\mu_i(\delta)$, $i = 1, \ldots, m$, (2.12)

3. The leading column coefficient matrix M_{hc} is unit upper triangular, (2.13)

4. The leading row coefficient matrix M_{hr} is the unit matrix. (2.14)

Thus a C1 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 [10]. Our definition of C1 canonical is almost the same up to a perturbation of the columns, so 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.

Once we have chosen the normalized solution of problem PA(δ) satisfying (2.10-14), then it will be a PA. 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

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

$$= p\omega - \sum_{i=1}^{p}\lambda_i(\delta) + \sum_{j=1}^{m}\mu_j(\delta) = p\omega.$$

which means that this solution of problem PA(δ) is a PA.

3. Solution of the scalar problem PA(δ)

Suppose we want to solve problem PA(δ) in the scalar case for $\delta = 0,1,\ldots,2\omega$ with the normalizations introduced before. The trivial solution for $\delta = 0$ 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 PA(δ), we shall treat them as being zero, which implies that $Y_0(z) = 0$.
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

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

If we agree to set $0 = O_-(z^{-\infty}) = O_+(z^{\infty})$, then we can say that it satisfies $P_0L_0^{-1} = O_-(z^{\omega+1})$. which is condition (1.6) but with $\delta = -1$.
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

$$[-1 \;\; F(z) \;\; -1]G_0 = 0. \tag{3.1}$$

Because both X_0 and Y_0 are $O_+(z^{\omega+1})$, it follows that (1.6) of problem PA(δ) is satisfied for 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 (1.6) for the entries of $G(z)$. We shall use this fact to generate polynomial matrices V_n, $n = 1,2,\ldots$ to 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

$$0 = \mu(0) = \mu(1) = \cdots = \mu(\delta_1 - 1) < \mu(\delta_1) = \cdots = \mu(\delta_2 - 1) < \mu(\delta_2) = \cdots .$$

I.e. the jumps in the denominator degree appear for the δ-values $\delta_1, \delta_2, \cdots$. The solutions of problem PA(δ) will be the same between two successive δ_i and 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

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

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

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

This is also true for $n = 0$ if we set $\partial N_{-1} = \omega + 1$ and $\partial M_{-1} = \mu(\delta_{-1}) = \mu(-1) = 0$.
Each V_n is made to satisfy $N_{0,n}(z)M_{0,n}(z)^{-1} = O_-(z^{\omega-\delta_n})$ 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 $\mu(\delta_{n-1})$. (3.2)
2. $M_{0,n-1}(z)^{-1}L_{0,n-1}(z) = O_-(z^{-1})$ (3.3)
3. $-P_{0,n-1}(z)$ is monic of degree $\omega - \mu(\delta_{n-1}) + 1$ (3.4)
4. $P_{0,n-1}(z)^{-1}N_{0,n-1}(z) = O_-(z^{-1})$. (3.5)
5. $N_{0,n-1}(z)M_{0,n-1}(z)^{-1} = O_-(z^{\omega-2\mu(\delta_{n-1})})$. (3.6)

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

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

with

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

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 $L_n = -Y_n^{-1}$. One can show that the degree of $M(z)$ is the minimal $\mu(\delta_n)$. Thus

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

L_n will make $-P(z)$ monic. $M(z)$ is monic and $M(z)^{-1}L(z) = O_-(z^{-1})$ by construction. So (3.2-4) are satisfied for $G(z)$. We next try to satisfy (3.5). 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 $M(z)$ but leave its degree untouched while $N_{0,n}(z)M_{0,n}(z)^{-1} = O_-(z^{-\gamma})$ with

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

Thus (3.6) will be satisfied automatically. Thus the matrix $V_n(z)$ is given by

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

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.

Example:

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

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

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

which gives

$$G_1 = \begin{bmatrix} 0 & z^7 \\ -1 & z^2-z+1 \\ -z^5-z^4-1 & z^4+z^2-z+1 \end{bmatrix}, \quad N_{0,1}M_{0,1}^{-1} = \frac{1-z+z^2+z^4}{1-z+z^2}.$$

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

so that

$$G_2 = \begin{bmatrix} -z^7 & z^8+z^7 \\ -z^2+z-1 & z^3 \\ -z^4-z^2+z-1 & z^3 \end{bmatrix}, \quad N_{0,2}M_{0,2}^{-1} = \frac{z^3}{z^3}.$$

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

and

$$G_3 = \begin{bmatrix} -z^8-z^7 & z^9+z^8-z^7 \\ -z^3 & z^4-z^2+z-1 \\ -z^3 & -z^2+z-1 \end{bmatrix}, \quad N_{0,3}M_{0,3}^{-1} = \frac{-1+z-z^2}{-1+z-z^2+z^4}.$$

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

$$G_4 = \begin{bmatrix} z^9+z^8-z^7 & z^{10}+2z^9+z^8 \\ z^4-z^2+z-1 & z^5+z^4-1 \\ -z^2+z-1 & -1 \end{bmatrix}, \quad N_{0,4}M_{0,4}^{-1} = \frac{-1}{-1+z^4+z^5}.$$

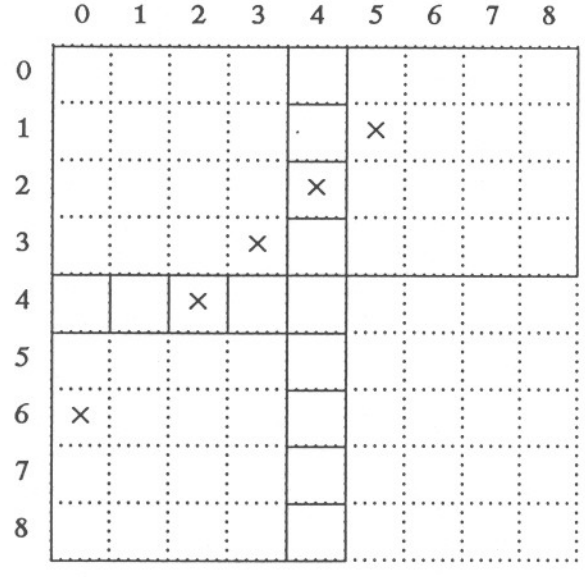
The last step gives

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

with

$$G_5 = \begin{bmatrix} z^{10}+2z^9+z^8 & z^{12}+z^{11}+z^7 \\ z^5+z^4-1 & z^7 \\ -1 & 0 \end{bmatrix}$$

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



4. Solution of the matrix problem PA(δ)

As for the scalar case, 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

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

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 (2.6) is satisfied for all $N_{0,n}$ and $M_{0,n}$ and therefore (2.6) 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.

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

Then because $|\underline{\mu}(\delta)| = |\underline{\lambda}(\delta)|$, the jumps for the $|\underline{\lambda}(\delta)|$ will occur at exactly the same places. This means that the solution of problem PA(δ) will not change between two successive δ_n.
To describe the construction of the V_n matrices that have to generate this change, we start from the assumptions, analog to (3.2-6), which $G_{0,n-1}$ should satisfy :

1. $M_{0,n-1}(z)$ is C1 canonical, with column 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 C2 canonical with column 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 C2 canonical is almost the same as C1 canonical, except that in the definition of C1 canonical we replace unit upper triangular by unit lower triangular. More precisely we call $-P(z)$ C2 canonical if it satisfies

1. The j-th column of $-P(z)$ has degree $\omega - \lambda_j(\delta) + 1$, $j = 1, \ldots, p$, (4.6)
2. The i-th row of $-P(z)$ has degree $\omega - \lambda_i(\delta) + 1$, $i = 1, \ldots, p$, (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)

It can be shown that by a rather complex algorithm it is possible to find 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 $N(z)$ form a solution of problem PA(δ) and such that

1. $M(z)$ is column reduced and $\partial \det M(z) = |\underline{\mu}(\delta_n)|$ (4.10)

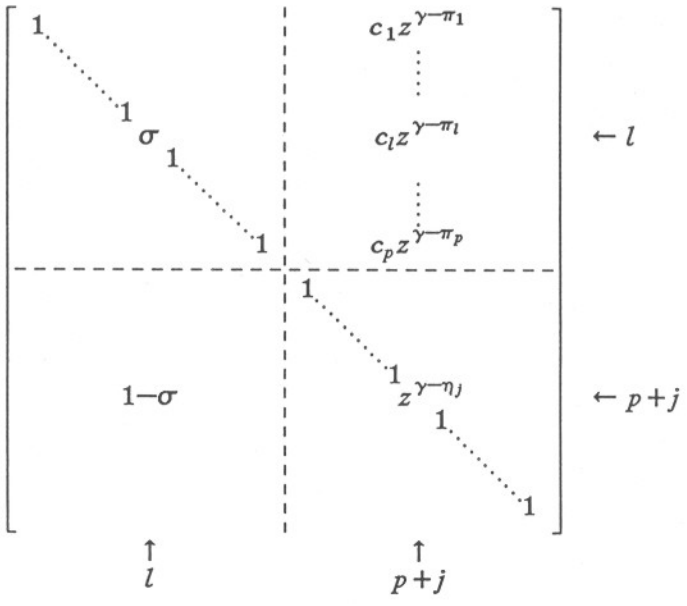
2. $P(z)$ is column reduced with column degrees of $P(z)$ and $L(z)$ satisfying

$$\partial P_j - \partial L_j > \omega - \delta. \tag{4.11}$$

To give an idea of what has to be done, we describe a general step in the construction of V_n^1. Suppose that the degree of the j-th column of $N(z)$ is η_j and that this violates (2.7). So we must eliminate its leading (column) coefficient. Because P_{hc} is nonsingular, we can write this leading coefficient as

$$P_{hc}[c_1\ c_2\ \cdots\ c_p]^T$$

Suppose the degrees of the columns of $P(z)$ are $\pi_1, \ldots, \pi_p$ and let π_l be (one of) the largest among these for which $c_l \neq 0$. Define $\gamma = \max\{\pi_l, \eta_j\}$ and $\sigma = 0$ if $\pi_l > \eta_j$ and $\sigma = 1$ if $\pi_l \leqslant \eta_j$. A multiplication with an elementary V-matrix of the form



will do the job. It eliminates the leading coefficient in column $p+j$ of G and replaces column l by the original column $p+j$ if $\sigma = 0$ i.e. if $\pi_l > \eta_j$. There will be no interchange if $\sigma = 1$ i.e. when $\eta_j \geqslant \pi_l$. Whether σ has to be 1 or 0 depends on the degrees of the columns involved. In any case, it is a unimodular transformation which respects (4.10-11). Moreover the column degrees of $M(z)$ after the application of V_n^1 will be a permutation of the elements of $\underline{\mu}(\delta_n)$. More details can be found in [12-16].

There may however be many solutions that can be generated in this way. By satisfying (4.1-5) we can make V_n unique. First we make $M(z)$ C1 canonical. This requires a multiplication with

$$V_n^2 = \begin{bmatrix} I_p & 0 \\ 0 & U_n^2 \end{bmatrix}$$

with U_n^2 unimodular.

Next we add polynomial multiples of $M(z)$ to $L(z)$ to satisfy (4.2). This is represented by a multiplication with

$$V_n^3 = \begin{bmatrix} I_p & 0 \\ Q_n^3 & I_m \end{bmatrix}$$

with Q_n^3 a polynomial matrix. This operation will not destroy condition (4.11).

Then $-P(z)$ is made C2 canonical by a matrix multiplication with

$$V_n^4 = \begin{bmatrix} U_n^4 & 0 \\ Q_n^4 & I_m \end{bmatrix}$$

with U_n^4 unimodular and Q_n^4 polynomial such that condition (4.2) is not destroyed.
The most complex step in the normalization is to satisfy (4.4). This requires the multiplication with

$$V_n^5 = \begin{bmatrix} I_p & Q_n^5 \\ 0 & U_n^5 \end{bmatrix}$$

with Q_n^5 polynomial and U_n^5 unimodular. The idea is that we take polynomial multiples of the first block column, correct this result with polynomial multiples of the second block column (of the type V_n^3) to make the degree of the $L(z)$-positions lower than the degree of the $M(z)$ without changing the leading term of the $P(z)$ positions and then add this result to the second block column to decrease the degrees of $N(z)$, while $M(z)$ remains C1 canonical. More details can be found in [16]. After these operations it can be shown that (4.5) is automatically satisfied so that we get

$$V_n(z) = V_n^1(z)V_n^2(z)V_n^3(z)V_n^4(z)V_n^5(z).$$

The algorithm for the matrix case is considerably more involved than the one for the scalar case. When applied to a problem with $p = m = 1$, it reduces to the one given in the previous section.

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 (2.7) 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

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

the following divisions for $n = 0,1,2,...$

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

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 chose 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 $M_n(z)^{-1}L_n = O_-(z^{-1})$.

It is clear from what we have said before that the algorithm presented in section 3 to solve problem PA(δ) is exactly the Euclidean algorithm applied to the polynomials $P(z) = -z^{\omega+1}$ and $N(z) = \sum_{k=0}^{\omega} F_k z^k$. Note that in general the gcd is obtained by adding one more step to the algorithm of section 3, i.e. for $\delta = 2\omega+1$.

In the example of section 3, the step for $\delta = 2\omega$ was already a solution for $\delta = 2\omega+1$, so the normalized gcd of P_0 and N_0 is -1, since they are indeed coprime.

The ideas of relating the Euclidean algorithm and Padé approximation has been given e. g. in [11].

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.

Suppose we have two polynomials $P(z) \in \mathbb{C}^{p\times p}$ and $N(z) \in \mathbb{C}^{p\times m}$. Then a greatest common left divisor (gcld) is defined as a polynomial matrix $R(z) \in \mathbb{C}^{p\times p}$ 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
$$R(z) = S(z)\overline{R}(z). \tag{6.2}$$

It can be shown [10] that 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.

Suppose $[P(z) \;\; N(z)]$ has full row rank and that

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

satisfies (4.1-4). (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

$$[P_0(z) \;\; N_0(z)] = [P(z) \;\; N(z)]\, V$$

is as required. These P_0 and N_0 matrices are then used as initializations instead of P and N. We can then apply the algorithm from section 4 and will find for $\delta = 2\omega+1$ (or for a smaller δ if it gives the same solution as for $\delta = 2\omega+1$) 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 C2 canonical polynomial matrix.

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

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

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 PA(δ) had

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

and thus are the PA-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 a general transformation 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

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

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. 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 PA's 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

$$\Sigma_\omega = \sum_0^\omega F_k z^k,\ L'_0 = 0_{m\times p}\ ,\ P'_0 = I_p\ .$$

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 als 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 PA. The left denominator is the transpose of a canonical matrix with row and column degrees $\lambda_i(\delta)$. Indeed, if (4.1,2,4,5) are satisfied with $k = n-1$ and $\delta = \delta_k$, then $-P_{0,k}^T$ is canonical with column degrees $\omega - \underline{\lambda}(\delta)+1$ iff $D_{0,k}^T$ is canonical with column degrees $\underline{\lambda}(\delta)$. This follows from the fact that

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

Since by (4.5)

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

and by (4.2)

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

we find that

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

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

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

So that

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

Because $-P_{0,k}$ is C2 canonical with row degrees $\omega-\lambda_i(\delta)+1$ it follows that $T_{1,k}$ is column reduced with column degrees $\lambda_j(\delta)$ and with $(T_{1,k})_{hc}$ the unit matrix. On the other hand we have from (4.11) that the column degrees of $P_{0,k}$ and $L_{0,k}$ are related by

$$\partial(P_{0,k})_j > \omega - \delta + \partial(L_{0,k})_j$$

and because

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

it follows from

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

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

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

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

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

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

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

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

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

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

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

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

Appendix : A more detailed description of the algorithm

Given a solution $G_{0,n-1}(z)$ of problem PA(δ) for $\delta = \delta_{n-1}$, we shall describe the different steps of the algorithm that will give a canonical solution for $\delta = \delta_n$. It will construct a unimodular matrix V_n such that $G_{0,n}(z) = G_{0,n-1}V_n$.

In this appendix we shall 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 $-\infty$.

An algorithm to find V_n^1 as in section 4

Suppose the i-th column of $N(z)$ is denoted by n_i.

Algorithm 1

1. $V_n^1 = I_{p+m}$
2. $\delta = \delta_{n-1}$
3. **While** $V_n^1 = I_{p+m}$
   - 3.1 $\delta = \delta+1$
   - 3.2 **For** $i = 1,2,\ldots,m$
     - 3.2.1 Set Hn_i = the coefficient of $z^{\omega-\delta+\mu_i}$ in n_i.
     - 3.2.2 **If** $Hn_i = 0$ **then** $V^{(i)} = I_{p+m}$.
     - 3.2.3 **If** $Hn_i \neq 0$ **then**
       - 3.2.3.1 write Hn_i as a linear combination of the columns of P_{hc}, i.e.
         $$Hn_i + P_{hc}[c_1,\ldots,c_p]^T = 0$$
       - 3.2.3.2 Choose some l such that
         $$\pi_l = \max\{\pi_j | c_j \neq 0\ ,\ j = 1,\ldots,p\}$$
         - 3.2.3.2.1 **If** $\pi_l > \eta_i$ **then**
           - 3.2.3.2.1.1 $V^{(i)} = E_{l,p+i}E_{p+i}(c_l)\prod_{\substack{j=1\\ j\neq l}}^{p} E_{j,p+i}(c_j z^{\pi_l-\pi_j})E_{l,p+i}(z^{\pi_l-\eta_i})$
           - 3.1.3.2.1.2 replace G by $GV^{(i)}$.
         - 3.2.3.2.2 **If** $\pi_l \leqslant \eta_i$ **then**
           - 3.2.3.2.2.1 $V^{(i)} = \prod_{j=1}^{p} E_{j,p+i}(c_j z^{\eta_i-\pi_j})$
           - 3.2.3.2.2.2 replace G by $GV^{(i)}$.
   - 3.3 Set $V_n^1 = \prod_{i=1}^{m} V^{(i)}$

An algorithm to make $M(z)$ C1 canonical

This algorithm will produce the matrix V_n^2 of section 4.
Suppose $M(z)$ is a column reduced $m \times m$ polynomial matrix. This matrix can be made canonical by an algorithm similar to the one described in [10,p481-486]. However, we shall prefer to give an other algorithm which uses the fact that we can start from a matrix M for which M_{hc} is already unit upper triangular. This is indeed what comes out of the previous algorithm. We can find a unimodular $U = U_n^2$ such that MU is C1 canonical. When applied to the whole matrix G, U is a block in the matrix $V = V_n^2 = \begin{bmatrix} I_p & 0 \\ 0 & U \end{bmatrix}$ which has to be initialized by I_{p+m}.

Algorithm 2

1. **For** $j = 1,2,\ldots,m$
   - 1.1 **For** $d = \mu_j, \mu_j - 1, \ldots, \mu_{\min}$
     - 1.1.1 **For** $i = m, m-1, \ldots, 1$
       - 1.1.1.1 **If** $i \neq j$ and $d \geq \mu_i$ and $M_{i,j,d} \neq 0$ **then**
         - 1.1.1.1.1 $\gamma = d - \mu_i$
         - 1.1.1.1.2 Replace $\begin{bmatrix} G \\ V \end{bmatrix}$ by $\begin{bmatrix} G \\ V \end{bmatrix} E_{p+i,p+j}(-M_{i,j,d} z^{\gamma})$

An algorithm to make $-P(z)_{hc}$ unit lower triangular

Before we shall compute the matrix V_n^3 of section 4, we shall first perform a step which will make it easier to construct V_n^4 lateron.
Suppose we know that $P(z)$ is column reduced and we want to find a unimodular matrix U to make $-P_{hc}$ unit lower triangular. Then U can be found as a block in the matrix $V = V_n^{4a} = \begin{bmatrix} U & 0 \\ 0 & I_m \end{bmatrix}$ which should be initialized by I_{p+m}.

Algorithm 4a

1. **For** $i = 1,2,\ldots,p$
   - 1.1 Choose some l such that $e = \pi_{il} = \min\{\pi_{ij} \mid \pi_{ij} = \pi_j \; j = i,\ldots,p\}$
   - 1.2 Set $q = -P_{i,l,e}^{-1}$
   - 1.3 Replace $\begin{bmatrix} G \\ V \end{bmatrix}$ by $\begin{bmatrix} G \\ V \end{bmatrix} E_{i,l} E_i(q)$
   - 1.4 **For** $j = i+1,\ldots,p$
     - 1.4.1 **If** $\pi_i \leqslant \pi_j$ and $P_{i,j,\pi_j} \neq 0$ **then**
       - 1.4.1.1 Set $\gamma = \pi_j - \pi_i$
       - 1.4.1.2 Replace $\begin{bmatrix} G \\ V \end{bmatrix}$ by $\begin{bmatrix} G \\ V \end{bmatrix} E_{i,j}(P_{i,j,\pi_j} z^{\gamma})$

An algorithm to find V_n^3 as in section 4

The problem is to find a polynomial matrix $Q = Q_n^3$, given $L(z) \in \mathbb{C}^{m\times p}[z]$ and $M(z) \in \mathbb{C}^{m\times m}[z]$, $M(z)$ being C1 canonical, such that

$M(z)^{-1}[L(z)+M(z)Q(z)] = O_-(z^{-1})$.

It is sufficient to make the row degrees of $L(z)+M(z)Q(z)$ less than the column degrees of $M(z)$.

The matrix $V = V_n^3 = \begin{bmatrix} I_p & 0 \\ Q & I_m \end{bmatrix}$ has to be initialized as I_{p+m}.

Algorithm 3

1. **For** $j = 1,2,\ldots,p$
   - 1.1 **For** $d = \alpha_j,\ldots,\mu_{\min}$
     - 1.1.1 **For** $i = m,m-1,\ldots,1$
       - 1.1.1.1 **If** $d \geqslant \mu_i$ and $L_{i,j,d} \neq 0$ **then**
         - 1.1.1.1.1 $\gamma = d - \mu_i$
         - 1.1.1.1.2 Replace $\begin{bmatrix} G \\ V \end{bmatrix}$ by $\begin{bmatrix} G \\ V \end{bmatrix} E_{p+i,j}(-L_{i,j,d}z^{\gamma})$

Algorithm to construct the matrix V_n^4 of section 4

We are now in a situation where M is already C1 canonical, $-P_{hc}$ is unit lower triangular and $M^{-1}L = O_-(z^{-1})$. We want to transform $-P$ into a C2 canonical matrix and keep the previous properties. We initialize $V = V_n^4$ as I_{p+m}.

Algorithm 4b

1. **For** $j = 1, \ldots, p$

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

      1.1.1 **For** $i = 1, \ldots, p$

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

            1.1.1.1.1 $\gamma = e - \pi_i$

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

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

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

               1.1.1.2.1.1 **If** $d \geq \mu_i$ and $L_{i,j,d} \neq 0$ **then**

                  1.1.1.2.1.1.1 $\gamma = d - \mu_i$

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

Algorithm to construct the matrix V_n^5 of section 4

The only thing left to do is to make $P^{-1}N = O_-(z^{-1})$. Therefore we shall make the row degrees of N less than the column degrees of P. We should however include corrections to keep M C1 canonical. The next algorithm will construct $V = V_n^5$ which has to be initialized as I_{p+m}.

Algorithm 5

1. **For** $j = 1, \ldots, m$

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

      1.1.1 **For** $i = 1, \ldots, p$

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

            1.1.1.1.1 $\gamma = e - \pi_i$

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

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

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

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

               1.1.2.1.1.1 $\gamma = d - \mu_i$

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