Matrix Padé Approximation: Definitions and Properties *

Guo-liang Xu †
Computing Center, Academia Sinica, Beijing, China
Adhemar Bultheel
Department of Computer Science, K.U.Leuven, Belgium

Abstract

Some proposals are made to give a general definition of matrix Padé approximants. Depending on the normalization of the denominator we define type I (constant term is the unit matrix) or type II (by conditions on the leading coefficient) approximants. Existence and uniqueness are considered, determinant expressions are given and relation among type I/II and left/right approximants are considered.

Contents

1 Introduction

The problem of classical (scalar) Padé approximation is well established and a number of textbooks appeared on the subject. E.g., the books from Baker and Graves-Morris [2] are among the more recent ones. However, when the given power series has matrix coefficients, it is not at all clear how the notion of Padé approximant should be generalized. In this paper we shall concentrate on possible definitions in the matrix case. To avoid a duplication of well known results, we shall suppose that the reader is sufficiently familiar with the problem of scalar Padé approximation.

There are already many papers that generalize the scalar Padé approximation problem in one way or another, but it is our conviction that the problem has never been considered in all its generality. There are e.g., many papers on the vector case, or on the square matrix case or even approximants in a general non-commutative algebra we studied. However, the non commutativity of the matrix multiplication is not so essential a complication, if this were the only one. Indeed some

*The work was done while the first author worked at the Department of Computer Science, K.U.Leuven
†Supported by the National Science Foundation of China for youth

papers did appear for the case of a square matrix function where the scalar normality condition immediately generalizes to the matrix case. Under this normality condition, all the recursive computations etc. from the scalar case do generalize (see e.g. [5, 6]). The more general problem of Padé approximation in a non commutative algebra has been considered by a lot of authors. A. Draux has contributions in this area and has compiled a commented bibliography of about 300 references on this topic [12]. However in the rectangular matrix case, the problem is more complicated since not only is $AB \neq BA$ in general, but it can be that BA does not even exist, even if AB does.

Also the vector case has attracted some attention in the literature. Different vector-Padé interpolants were introduced by P.R. Graves-Morris (see e.g., [14]). The problem here is to approximate a number of functions by rational functions with a common denominator. Because the denominator is a scalar polynomial, the problem is relatively simple. There is e.g., no problem in its normalization or in pinning down its degree. In the general case, the denominator is a square matrix polynomial for which the problem of defining a normalization or defining its degree leaves a lot of possibilities.

There is also an extensive literature in linear system theory on the problem of minimal partial realization which is a problem that in a certain sense is equivalent with Padé approximation (see e.g. [8]). If the system has many inputs and many outputs, then one gets a matrix approximation problem which is closely related to a matrix Padé problem. The main difference is that the given power series there is a series in z^{-1} while the numerator and denominator of the rational approximant are polynomials in z. Moreover, minimality is an important issue in those applications. In the matrix case, this means that the degree of the determinant of the denominator has to be minimal. An effort has been done to translate these results into Padé results (see [7]). Therefore a reformulation of the Padé problem as a minimal Padé problem is given, and it is the latter that is generalized to the matrix case. In this text we shall refer to it as the problem mPA or MmPA in the matrix case.

In a number of papers, the first author has considered a general matrix Padé problem for a rectangular matrix series (see e.g., [20],[21]). It appeared that the solutions of the MmPA problem were matrix Padé approximants that were not covered by this theory. This motivated a more general approach to the problem of defining matrix Padé approximants.

2 The Definition of Matrix Padé Approximation

Let $f(z)$ be a given power series with matrix coefficients, i.e.,

$$f(z) = \sum_{k=0}^{\infty} c_k z^k, \qquad c_k \in \mathbf{C}^{p\times m}, \tag{2.1}$$

where $\mathbf{C}^{p\times m}$ consists of all $p \times m$ matrices with their elements in the complex plane $\mathbf{C}$. We want to determine a (right) rational approximating function $N(z)M(z)^{-1}$ to f by the following equation:

$$f(z)M(z) - N(z) = R(z). \tag{2.2}$$

Where $N(z)$ and $M(z)$ are $p \times m$ and $m \times m$ matrices respectively with polynomials as their elements. We call $R(z)$ the residual of the approximant. We shall impose some lower bounds on the orders of the entries in $R(z)$ and some upper bounds on the degrees of the entries in $M(z)$ and $N(z)$.

This seems to be a fair proposal to generalize the definition of Padé approximation (PA) to the matrix case. However, depending on the purposes which one has in mind, this leaves us with still many ways of defining the matrix Padé approximants (MPAs) as we shall see later. Whatever the definition is, since we refer to them as matrix Padé approximants, the MPAs should have some relations or connections with classical (i.e., scalar) PAs. Therefore, there are some principles in defining the approximants which we should follow. Some of these are listed below.

(i) (matrix) MPAs should be a generalization of (scalar) PAs. This means that when $p = m = 1$ the matrix Padé approximants should coincide with the scalar Padé approximants. The difficulty is that one can give many formulations of the Padé problem which are all equivalent in the scalar case, but give completely different descriptions when they are generalized to the matrix problem.

(ii) Solvability. In general the defining equations should have a unique solution. By this we mean that the number of unknowns i.e., the number of free parameters to be determined in the approximants, is equal to the number of approximation conditions, i.e., the number of equations which determine these unknowns. We say "in general" since it may happen that in some cases the problem has none or many solutions. It is based on this condition that one can give in the scalar case upper bounds for the degrees of numerator and denominator and then their sum has to be a strict lower bound for the order of approximation. More generally, one can give bounds for two of these three numbers and then a bound for the third one follows. This condition is commonly accepted by those who generalized the scalar Padé approximants. In the vector case, this idea has been used to fix a degree for the scalar denominator and e.g., one global order of approximation. The numerator degrees are then bounded to satisfy the solvability condition. Also the converse, where numerator and denominator degrees are bounded, can be done. Lower bounds for the orders of approximation per component then follow from the same conditions again.

(iii) The MPAs thus defined should have some nice properties just as scalar PAs have. E.g., invariance under linear fractional transformation (see section 6).

In recent years many extentions and generalizations of the scalar Padé problem have been established. Matrix Padé approximants were, in our opinion, never treated in a sufficiently general form. So we can ask why there has not been a good definition of MPAs so far. What makes this problem so much more complicated than its scalar counterpart and what can we do about it ? We give some possible reasons below and we shall try to propose a solution.

(a) **Non-commutativity :** The multiplication of matrices is not commutative in general.

(b) **Choice of degrees and orders :** There are many possibilities in choosing the degrees of the numerator and denominator and the order of the residual.

(c) **Normalization :** There are many ways of normalizing a solution to the problem. We shall consider two possibilities and refer to them as type I and type II approximants.

(a) **Non-commutativity :** As a consequence of this non-commutative multiplication we have to define right MPAs and left MPAs separately. The relation (2.2) is for a right approximant. A left approximant $\tilde{M}(z)^{-1}\tilde{N}(z)$ would, instead of (2.2), satisfy an equation like

$$\tilde{M}(z)f(z) - \tilde{N}(z) = \tilde{R}(z)$$

where now $\tilde{N}(z)$ and $\tilde{M}(z)$ are $p \times m$ and $p \times p$ polynomial matrices respectively.

(b) **Choice of degrees and orders :** Concerning this point we try to start with the most general situation, in which we choose a degree for each entry of the numerator and a degree for each entry of the denominator. Also with each entry of the residual we associate an order. It is by this general approach that we can obtain considerable extensions of most known matrix Padé approximants. Unfortunately, if we want to exploit this generality, we shall have to deal with much more complicated conditions, normalizations, algorithms etc. Rather than polynomials with matrix coefficients, we consider matrices with polynomial entries. To this purpose, we introduce the following notations:

$$\mathbf{H}_k := \{p(z) : p(z) = \sum_{i=0}^{k} a_i z^i, \quad a_i \in \mathbf{C}\},$$

and

$$\mathbf{E}_k := \{e(z) : e(z) = \sum_{i=k+1}^{\infty} a_i z^i, \quad a_i \in \mathbf{C}\}.$$

The nonnegative integers are denoted by $\mathbf{Z}_+$ and $\mathbf{Z}_+^{p\times m}$ denotes all the $p\times m$ matrices with elements in $\mathbf{Z}_+$. For $V=(v_{ij})\in\mathbf{Z}_+^{p\times m}$, we introduce the notations

$$\mathbf{H}_V^{p\times m}=\{P(z)=(p_{ij}(z))_{i,j=1}^{p,m}:p_{ij}(z)\in\mathbf{H}_{v_{ij}}\},$$

and

$$\mathbf{E}_V^{p\times m}=\{R(z)=(r_{ij}(z))_{i,j=1}^{p,m}:r_{ij}(z)\in\mathbf{E}_{v_{ij}}\}.$$

Sometimes we shall express the fact that $R(z)\in\mathbf{E}_V^{p\times m}$ as $R(z)=O(z^{V+1})$. In this case we consider $P(z)\in\mathbf{H}_V^{p\times m}$ as a table of $p\times m$ polynomials over $\mathbf{C}$. Each entry has its own degree. We could also consider it as a column of p polynomials over $\mathbf{C}^{1\times m}$. In that case, each row has its degree. Thus if $V=(v_i)_{i=1}^p\in\mathbf{Z}_+^{p\times 1}$, then the notation $\mathbf{H}_V^{p\times m}$ denotes

$$\mathbf{H}_V^{p\times m}=\{P(z)=(p_i(z))_{i=1}^p:p_i(z)\in\mathbf{H}_{v_i}^{1\times m}\}.$$

The v_j correspond to row degrees. Similarly, we can consider these polynomial matrices as a row of vector polynomials. If $V=(v_j)_{j=1}^m\in\mathbf{Z}_+^{1\times m}$, then the notation $\mathbf{H}_V^{p\times m}$ denotes

$$\mathbf{H}_V^{p\times m}=\{P(z)=(p_j(z))_{j=1}^m:p_j(z)\in\mathbf{H}_{v_j}^{p\times 1}\}.$$

Now the v_j correspond to column degrees. Finally, if the polynomial matrix is considered as a matrix polynomial, i.e., a polynomial with matrix coefficient, it has a scalar degree. The set of all the matrix polynomials of degree at most $V\in\mathbf{Z}_+$ is denoted as $\mathbf{H}_V^{p\times m}$. It is of course trivial to identify these sets by adding or deleting zero coefficients at appropriate places and we shall not distinguish between them. It will thus depend on the dimension of V whether we mean element degrees, row degrees, column degrees or matrix degrees. We shall refer to the latter as a uniform degree because it is the same for all elements. For example, for $V=(v_i)\in\mathbf{Z}_+^{p\times 1}$ we make the identification $\mathbf{H}_V^{p\times m}=\mathbf{H}_{\tilde V}^{p\times m}$ with $\tilde V=(\tilde v_{ij})$ if $\tilde v_{ij}=v_i$ for $j=1,\ldots,m$. This identification of the two sets means that we do not distinguish between V and $\tilde V$. Similar conventions will also be used in relation with the orders of the entries of the elements of $\mathbf{E}_V^{p\times m}$.

Having the previous conventions in mind, we define additions and subtractions of degree-matrices or order-matrices of unequal dimensions in a natural way. (Remember the identification of $V\in\mathbf{Z}_+^{p\times 1}$ and $\tilde V\in\mathbf{Z}_+^{p\times m}$.) Two examples are

(i) If $V\in\mathbf{Z}_+^{s\times t}$ and $W\in\mathbf{Z}_+$, then $V+W=(v_{ij}+W)_{i,j=1}^{s,t}$.
(ii) If $V\in\mathbf{Z}_+^{s\times 1}$ and $W\in\mathbf{Z}_+^{1\times t}$, then $V+W=(v_i+w_j)_{i,j=1}^{s,t}$.

It may be somewhat confusing to add matrices of different dimensions but if you just remember that a row matrix should be expanded to a full matrix of appropriate dimension with every row duplicating the given one, a column matrix should be expanded to a full matrix by duplicating this column and a scalar is expanded to a full matrix with all identical elements, then this operation is not so unusual. An expression $W-V=U$ with W and V in $\mathbf{Z}_+^{s\times t}$ and $U\in\mathbf{Z}_+$ is a short way of telling you that the difference between the corresponding elements in W and V is the constant value U for each entry. Along the same line a relation $V\geq 0$ for some $V\in\mathbf{Z}_+^{s\times t}$ has to be understood on an elementwise basis. It means that all the entries of V are nonnegative.

(c) **Normalization :** As a normalization in the scalar case, one usually chooses the denominator to be monic or comonic. Such a normalization serves two purposes. It fixes a unique numerator-denominator pair for a given rational Padé approximant and it fixes a coefficient so that the true parameters of the problem which can be used to satisfy the approximating conditions show clearly. A rational function of type v/u has $v+u+2$ coefficients but only $v+u+1$ parameters are available to satisfy the interpolation conditions and one coefficient is pinned down by the normalization. Thus a proper question to ask in the matrix case is : can we find a unique numerator-denominator pair $(\tilde N(z),\tilde M(z))\in\mathbf{H}_{\tilde V}^{p\times m}\times\mathbf{H}_{\tilde U}^{m\times m}$ for a certain approximant $R(z)$ which is right coprime and which is uniquely defined by $R(z)$. If $R(z) = N(z)M(z)^{-1}$ with $(N(z), M(z)) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m}$ right coprime, is one possible representaion of $R(z)$, can we then find some G such that $(\tilde{N}(z), \tilde{M}(z)) = (N(z)G, M(z)G)$ satisfies the requirements given above ? Since the greatest common right divisor is only defined up to a right unimodular factor (i.e., a polynomial matrix G with $\det G \in \mathbf{C}\backslash\{0\}$), we should allow a unimodular G in general. Note however that the normalization will change the type of the polynomial pair, i.e., $U \neq \tilde{U}$ and $V \neq \tilde{V}$ in general, even in the case where G is a constant matrix. Here we are confronted with another nasty complication as compared to the scalar case. If in the scalar case a coprime solution pair exists, then we could always find a normalized one with the same degree structure. For the matrix case, it depends on the V and U matrices how many coefficients can be fixed without changing neither the degree type, nor the rational function. Such a problem is very difficult and only decidable in simple cases like e.g., $U \in \mathbf{Z}_+$. We can then take $M(0) = I$ or make $M(z)$ monic i.e., $M(z) = z^U I + \cdots$. A way around this problem is to just impose some normalization and see to it that the number of equations equals the number of free parameters left. We can for example take

$$M(0) = I. \tag{2.3}$$

If $M(z)$ is an element from $\mathbf{H}_U^{m\times m}$ for some $U = (u_{ij}) \in \mathbf{Z}_+^{m\times m}$, then, under the normalizing condition (2.3), the j-th column of $M(z)$ has $\sum_{k=1}^m u_{kj}$ unkowns to be determined by approximation conditions. The normalization (2.3) is of course just an example and there are many other possibilities. It is however a quite natural condition since it guarantees that $N(z)M(z)^{-1}$ has no singularities at $z = 0$ while $R(z)$ should approximate the given function precisely at the origin. A "monic" normalization of $M(z)$ is not so simple for general U since it is not very clear what should be the leading coefficient. However, it is possible to think of a suitable substitute. We shall not do this for the type I MPA definition given below where we use the comonic normalization (2.3).

For the type II MPA definition we do have something in the style of such a monic normalization. Its origin is however rather different. That is why we shall use it for another type of MPAs. In problems of minimal partial realization, one tries to fit a number of Markov parameters by a rational transfer function with a minimal complexity. This means that the degree of the determinant of the denominator should be as low as possible while the precise degrees of the entries are not an important issue. So the problem arises of representing the denominator of the rational function in its simplest possible form. To this problem there exists a solution which is a well known procedure in linear system theory. As the reader will know, a constant matrix can be brought into an echelon form, which is its simplest form to which it can be reduced using only elementary column transformations. For a polynomial matrix, something similar is possible where now elementary unimodular transformations are allowed. Note that these can change the type U of $M(z)$, but leave the degree of its determinant unchanged. It can be shown (see [15]) that by using unimodular transformations, any square matrix polynomial can be brought into a form such that its row degrees are equal to its column degrees. The leading row (or column) coefficient matrix M^{hr} (or M^{hc}) is then the matrix formed by the leading row (or column) coefficients. The second method for normalization we shall use in this paper is to impose conditions on M^{hc} and M^{hr}. We shall then say that $M(z)$ is brought into a canonical form. Two variants ($C1$ or $C2$ canonical forms) are defined below (see[7]).

Definition 2.1 ($C1/2$ canonical) *We say that $M(z) \in \mathbf{H}_V^{m\times m}$ is $C1$ (or $C2$) canonical iff*

(i) There is a vector $U \in \mathbf{Z}_+^{m\times 1}$ such that $M(z) \in \mathbf{H}_U^{m\times m} \cap \mathbf{H}_{U^T}^{m\times m}$. This means that row and column degrees are the same.

(ii) The leading row (or column) coefficient matrix M^{hr} (or M^{hc}) of $M(z)$ is the unit matrix.

(iii) The leading column (or row) coefficient matrix M^{hc} (or M^{hr}) of $M(z)$ is unit upper (or lower) triangular.

A polynomial matrix for which M^{hc} (M^{hr}) is regular is called *column (row) reduced*. This means that the above canonical forms are column as well as row reduced. Consequently, the degree of the determinant is equal to the sum of the column degrees and equal to the sum of the row degrees, which is $\sum_{i=1}^{m} u_i$. Note that in the scalar case $C1$ and $C2$ canonical means the same thing viz., $M(z)$ is monic. Note also that if $M(z)$ is $C1$ canonical with degree $U = (u_i)_{i=1}^{m}$, then there are

$$\sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1$$

unknowns for determining the j-th column of $M(z)$.

As we have said before, the previous normalization is inspired by the minimal partial realization problem which is closely related to, but not exacly the same as, the problem of Padé approximation. This relation is generally known in the scalar case (see e.g., [13]). Results from one field have been translated and applied to the other one, algorithms have been adapted and exchanged, etc. but in our opinion there are basic differences that may not be very important for the scalar case, but which do matter in the matrix case. One of them is the notion of minimality of the approximant and another one the approximation being at $z = \infty$ rather than in the origin. In the matrix case, probably because of the lack of a good definition of matrix Padé approximants, correspondences were mentioned but never explained. The report [7] by A. Bultheel and M. Van Barel and the Ph. D. thesis [17] seem to be exceptions to this. Based on a number of reports on the minimal partial realization problem (see [16]), the authors have in [7] defined a kind of matrix Padé approximants which they called (scalar) minimal Padé Approximants (mPAs) and Matrix minimal Padé Approximants (MmPAs) because they have minimality properties in the degrees of numerator and denominator. In our opinion, this is the true Padé-like problem that really corresponds to the minimal partial realization problem and is a somewhat different approach to Padé approximation which is quite interesting and as we feel it, sometimes a more natural problem to solve. Because it is a non standard definition, and because we shall need it in the sequel, we repeat the definition of the MmPA problem below. For a rather extensive study of this problem see [7] and [17]. In the minimal partial realization problem, one starts with the coefficients $c_1, \ldots, c_\omega$ of the series $f(z) = \sum_{k=1}^{\infty} c_k z^{-k}$ and the problem is to find polynomials $N(z)$ and $M(z)$ of appropriate dimensions in the variable z, such that $f(z) - N(z)M(z)^{-1} = O(z^{-\omega-1}), \quad z \to \infty$ and such that the degree of det $M(z)$ is minimal. As we know by now, we can always choose $M(z)$ to be canonical. The approximation condition implies that $N(z)M(z)^{-1}$ is strictly proper. This, and the regularity of M^{hc} and M^{hr} imply that there is no problem in expanding $N(z)M(z)^{-1}$ around infinity. This is *not* turned into a Padé problem by simply replacing z by z^{-1}. Instead, the minimal partial realization problem can be easily translated into the following problem. (see [7] and [17] for all the details.)

Definition 2.2 (MmPA Problem) *Given $f(z) = \sum_0^{\infty} c_k z^k$, some nonnegative integer ω and an integer $\delta \in [-\omega, \omega]$, find polynomials $M(z)$ and $N(z)$ such that*

1. $f(z)M(z) - N(z) = O(z^{\omega+1}), \qquad z \to 0$
2. $\eta_j \le \mu_j - \delta$ *for* $j = 1, \ldots, m$. *Here* η_j *and* μ_j *are the degrees of the* j*-th column of* $N(z)$ *and* $M(z)$ *respectively.*
3. $\sum_{j=1}^{m} \mu_j$ *as low as possible.*
4. $M(z)$ *is* $C1$ *canonical.*

It follows from the above definition that the μ_j are the Kronecker indices for the sequence $\{c_\omega, c_{\omega-1}, \ldots, c_{-\delta+1}\}$ (see definition 5.1). It can also be shown that the solution to the MmPA problem as it is formulated above may not be unique. It can be made unique by imposing the extra condition that

5. $\nu_i \le \omega - \lambda_i \quad for \quad i = 1, \ldots, p$

where the ν_i are the row degrees of $N(z)$ and the λ_i are the dual Kronecker indices for the same sequence as mentioned above. With this extra condition, the solution is called canonical and this is the definition used in [7].

One can see from the definition of the MmPA problem that, even in the scalar case it is different from the PA problem. But mPA does have a close relationship with PA. In fact, any solution of the scalar mPA problem is one of the PAs in the Padé table (see [7]). It is one of the main advantages of considering problem mPA instead of problem PA that we can define a table of minimal Padé approximants which is parametrized in only two scalar parameters ω and δ, *even in the matrix case.* In the scalar case, this ω-δ–table is essentially the same as the classical Padé table (see [10]). As a matter of fact, in the scalar case, one can consider the minimal Padé approximation problem as a problem where one looks in the Padé table among all PAs of a certain approximation order ω for the one which is the simplest. If one considers the maximum of the degree of numerator and denominator as a measure of complexity, then one gets a PA which is near the main diagonal of the Padé table. To get other approximants, one can impose a relative importance of numerator and denominator degree by playing with the parameter δ.

In the definition given above, we fixed the order ω and used δ to parametrize the difference in the degrees. We could as well have chosen to fix the numerator degree and use δ to parametrize the order of approximation. The latter approach was used in the vector case in [9]. The same ideas could also be used for more general rational interpolation. See [18].

Therefore we shall also try to define Padé approximants in the matrix case such that they are solutions of an MmPA problem. The last problem setting was the one that partly motivated this work.

Now we are ready to give the following definitions of type I and type II right MPAs. Type I refers to the normalization (2.3) and type II refers to a $C1$ canonical normalization.

Definition 2.3 (Type I MPA) *Let $f(z) \in \mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, $U = (u_{ij}) \in \mathbf{Z}_+^{m\times m}$ and $W = (w_{ij}) \in \mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{k=1}^{m} u_{kj}, \quad j = 1, 2, \ldots, m, \tag{2.4}$$

and $W - V \geq 0$. Then the right MPA problem of the first type is denoted by ${}^R(V, U, W; f)_I$ and consists in finding polynomials $N(z) \in \mathbf{H}_V^{p\times m}$ and $M(z) \in \mathbf{H}_U^{m\times m}$ such that

1. *$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}$.*
2. *$M(0) = I$.*

The collection of all rational functions $N(z)M(z)^{-1}$ formed by the solutions of ${}^R(V, U, W; f)_I$ is denoted by ${}^R[V, U, W; f]_I$.

Note that the matrices V, U and W are completely general in this definition. In the next definition, the U matrix which indicates the denominator degree will be a vector, as motivated when we introduced the canonical normalizations. The V and W matrices are still general.

Definition 2.4 (Type II MPA) *Let $f(z) \in \mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, $U = (u_j) \in \mathbf{Z}_+^{m\times 1}$ and $W = (w_{ij}) \in \mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{\substack{u_i > u_j \\ 1 \leq i \leq m}} u_j + \sum_{\substack{u_i \leq u_j \\ 1 \leq i \leq m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \leq i < j}} 1, \quad j = 1, 2, \ldots, m, \tag{2.5}$$

and $W - V \geq 0$. Then the right MPA problem of the second type is denoted by ${}^R(V, U, W; f)_{II}$ and consists in finding polynomials $N(z) \in \mathbf{H}_V^{p\times m}$ and $M(z) \in \mathbf{H}_U^{m\times m}$ such that

1. $f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}$.

2. $M(z)$ *is* $C1$ *canonical with degree* U.

The solution set of ${}^R(V,U,W;f)_{II}$ is denoted by ${}^R[V,U,W;f]_{II}$.

Conditions (2.4) and (2.5) make the number of (scalar) coefficients to be found equal to the numbers of (scalar) equations. Therefore we shall call them *solvability equalities.*

Besides these two possibilities for the normalization of the denominator, there are of course many other choices that could be made. Each of them can give another type of Padé approximant and it is not sure, as we shall show for the type I and type II approximants that the different normalizations are equivalent in all situations. They give in many cases completely different approximants.

There are some special cases of the above definitions which we should mention here, because each of them can be regarded as a different definition of MPA. We shall make different choices for the degrees and orders and see what (dis)advantages this may have. Before we do this, we introduce the following lemma.

Lemma 2.1 *Let* $U \in \mathbf{Z}_+^{m\times 1}$. *Then*

$$\sum_{i=1}^m u_i = \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1\le i\le m}} u_i + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j=1,2,\ldots,m, \tag{2.6}$$

if and only if

$$\begin{cases} |u_i - u_j| \le 1, & for\ \ \forall i,j \\ u_1 \ge u_2 \ge \cdots \ge u_m. \end{cases} \tag{2.7}$$

Proof. If we bring the middle term of the right hand side of (2.6) to the left of the equality sign, we see that (2.6) holds if and only if

$$\sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_i = \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j=1,2,\ldots,m.$$

Now we bring the first term of the right to the left of the equality sign to get

$$\sum_{\substack{u_i > u_j \\ 1\le i\le m}} (u_i - u_j) = \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j=1,2,\ldots,m.$$

If we split up the left hand side as follows

$$\sum_{\substack{u_i > u_j \\ 1\le i< j}} (u_i - u_j) + \sum_{\substack{u_i > u_j \\ j\le i\le m}} (u_i - u_j) = \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \qquad j=1,2,\ldots,m,$$

then it follows that

$$u_i - u_j = 1, \quad \text{for } \ u_i > u_j,\ 1\le i<j, \ \text{ and } \ u_i - u_j \le 0 \quad \text{for } \ j\le i\le m.$$

This is (2.7). □

Note that (2.7) means that all the numbers u_i are equal to each other except for a possible drop by 1 at a certain moment, after which all the numbers stay at this value.

Now we list some special cases.

D1. V, U and $W \in \mathbf{Z}_+$.

This is the case where numerator and denominator are considered as polynomials with matix coefficients, i.e., as matrix polynomials. If $p = m$, this is the most popular case. Most of the literature on matrix Padé approximation in the past decades is about this situation. It doesn't matter so much that the coefficients are matrices. The main difficulty is that their product is not commutative anymore and that zero divisors can exist. That is the reason why the Padé approximation problem allows a rather complete analysis. The many publications that fall under this classifications need not be repeated here. An excellent survey of the literature can be found in [12]. The matrix case for $p \neq m$ was discussed in [20]. In this case, both (2.4) and (2.5) lead to the same solvability equality

$$pw = pv + mu.$$

This means that p should be a divisor of mu. This may be a severe limitation in practice. Most of the papers dealing with the matrix case explicitly and not just with a non commutative algebra do have such divisibility conditions.

D2. $V \in \mathbf{Z}_+^{1\times m}$, $U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{1\times m}$.

This means that we do the Padé problem column-wise. In this case, (2.4) for the type I approximants becomes

$$pw_j = pv_j + mu_j, \quad j = 1, 2, \ldots, m.$$

This implies that p should divide mu_j for every j. This is an even stronger limitation than we already had in the previous case. For type II approximants, condition (2.5) leads to

$$pw_j = pv_j + \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

This means that the total number represented by the last three terms of the right hand side should be divisible by p. Again a condition that is not simple to deal with.

D3. $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{1\times m}$ and $W \in \mathbf{Z}_+^{p\times 1}$.

Checking the type I condition (2.4) we get

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k + mu_j, \quad j = 1, 2, \ldots, m.$$

This implies that

$$u_1 = u_2 = \cdots = u_m. \tag{2.8}$$

Thus we get a rather simple condition. There is a global degree for the denominator while numerator degrees and orders are considered row-wise. For the type II approximants, this situation is rather restrictive since equation (2.5) leads to (2.7) as derived in Lemma 2.1.

D4. $V \in \mathbf{Z}_+^{1\times m}$, $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+^{1\times m}$.

This is kind of complementary to the situation of D3. The solvability equality (2.4) for type I approximants becomes

$$pw_j = pv_j + \sum_{i=1}^{m} u_i, \quad j = 1, 2, \ldots, m.$$

This means that $\sum_{i=1}^{m} u_i$ should be divisible by p and $W - V \in \mathbf{Z}_+$, i.e., $w_j - v_j$ does not depend on j. The latter condition is like in the MmPA problem but the divisibility condition is an extra. For type II approximants (2.5) leads to the same situation as for D2.

We can go on in this manner considering row and column degrees and orders. There are eight possibilities in all. We skip a few and give one more interesting case that we shall see reappear in the sequel.

D5. $V \in \mathbf{Z}_+^{p\times 1}$, $U \in \mathbf{Z}_+^{m\times 1}$ and $W \in \mathbf{Z}_+^{m\times 1}$.
Now (2.4) becomes

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k + \sum_{k=1}^{m} u_k, \tag{2.9}$$

which means that type I MPAs are always possible, at least they can always satisfy the solvability equality without extra restriction. That is why this case is important in this paper. Condition (2.5) for type II MPAs becomes

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k + \sum_{\substack{u_i > u_j \\ 1 \le i \le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1 \le i \le m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \le i < j}} 1, \quad j = 1, 2, \ldots, m.$$

This holds only for U satisfying (2.7).

We picked some possibilities for row or column degrees and orders merely as an illustration of the type of conditions one gets for solvability equalities. Many other choices are possible to select degrees and orders. We never claim that a particular choice is the best one to make. As we said before almost all the MPAs considered in the literature fall into the class D1. A particular case which recieved more attention was the vector case where $m = 1$. We then get the cases D2 and D6 which turn out to be the same, since there is a scalar denominator so that $U = u \in \mathbf{Z}_+$ while $V \in \mathbf{Z}_+^{p\times 1}$ and $W \in \mathbf{Z}_+^{p\times 1}$. Things then become much more simple because the solvability equality is for both type I and type II MPAs just

$$\sum_{k=1}^{p} w_k = \sum_{k=1}^{p} v_k + u.$$

This condition gives only a constraint for the sums involved. Within this frame, still several options for a definition can be taken. Different choices of how the numerator degrees or approximation orders are distributed over the components can be made. One can consult the work of Graves-Morris [14] or van Iseghem [19] and the references therein.

Of course, everything that has been said so far on right MPAs can now be adapted for left MPAs. In the notation we use L instead of R. Writing out all the details is an easy task which we gladly leave to the reader. We should only mention that, in the definition of the second type left MPA, the denominator should be $C2$ canonical instead of $C1$ canonical.

Whenever possible we simplify the notation if it gives no ambiguity. Thus the indication R or f or I or II shall be dropped if it is clear from the context what approximant is being meant.

3 The Existence of Matrix PAs

It is well known that in the scalar case the PA problem leads to a linear system of equations, which are called Padé equations. The numerator coefficients can be found from the data and the denominator coefficients. Therefore, the basic problem is to set up the system of linear equations which defines the denominator coefficients. The existence of the solution of this system implies the existence of a Padé approximant. There is an analogue to this in the matrix case. Setting up these equations in the matrix case is not a difficult job to do. Existence theorems will merely express that these systems have a solution. For the general case it is rather difficult to isolate certain classes of functions for which all the MPAs exist, even for MPAs of a certain type. The Stieltjes series, which are well known in the scalar case to guarantee the existence of all PAs, have been generalized to the square matrix case by Basu and Bose in [3] but only for matrix degrees and orders, i.e., MPAs of type D1. Without further specification of the approximants, this is impossible to do in general. Thus the conditions stand as they are. They are however useful because they introduce a notation that shall be used in the next section to give determinant expressions for the approximants. Also they will show incidentally the special structure of the systems which can be exploited to design fast algorithms for computing the approximant. In many cases the matrices of the systems have a low displacement rank so that they can be solved by specially designed algorithms. These computational aspects will be treated in a separate publication [22].

In order to express the problems ${}^R(V,U,W;f)_I$ and ${}^R(V,U,W;f)_{II}$ as systems of linear equations, we introduce the following notations. For

$$g(z)=\sum_{k=0}^{\infty} a_k z^k, \quad a_k \in \mathbf{C}^{s\times t},$$

let

$$T^l_{mn}(g) = \begin{bmatrix} a_l & a_{l-1} & \cdots & a_{l-n+1} \\ a_{l+1} & a_l & \cdots & a_{l-n+2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{l+m-1} & a_{l+m-2} & \cdots & a_{l+m-n} \end{bmatrix} \in \mathbf{C}^{ms\times nt}$$

be a block Toeplitz matrix. We adopt the usual convention that $a_k = 0$ for $k < 0$. If $s = t = 1$, then T^l_{mn} is an ordinary Toeplitz matrix. Let $f = (f_{ij}) \in \mathbf{C}^{p\times m}[[z]]$ be given. For the problem ${}^R(V,U,W;f)_I$ where we suppose V, U and W fixed, we introduce the matrices

$$E = (e_{ij}) = W - V \tag{3.10}$$

and

$${}^RH^j_I = \begin{bmatrix} T^{v_{1j}}_{e_{1j},u_{1j}}(f_{11}) & T^{v_{1j}}_{e_{1j},u_{2j}}(f_{12}) & \cdots & T^{v_{1j}}_{e_{1j},u_{mj}}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_{pj}}_{e_{pj},u_{1j}}(f_{p1}) & T^{v_{pj}}_{e_{pj},u_{2j}}(f_{p2}) & \cdots & T^{v_{pj}}_{e_{pj},u_{mj}}(f_{pm}) \end{bmatrix}, \tag{3.11}$$

and

$${}^RB^j_I = \begin{bmatrix} T^{v_{1j}+1}_{e_{1j},1}(f_{1j}) \\ T^{v_{2j}+1}_{e_{2j},1}(f_{2j}) \\ \vdots \\ T^{v_{pj}+1}_{e_{pj},1}(f_{pj}) \end{bmatrix}. \tag{3.12}$$

Using these notations, the problem of determining the j-th column of the denominator $M(z)$ in ${}^R(V,U,W;f)_I$ can be expressed by the following equations.

$${}^RH^j_I\ X = -\ {}^RB^j_I. \tag{3.13}$$

If the j-th column $M_j(z)$ of $M(z)$ is determined, the j-th column $N_j(z)$ of the numerator $N(z)$ can be easily found from

$$N_j(z) = (fM_j)^{(V_j)}(z) = f(z)M_j(z) \bmod z^{V_j},$$

where V_j is the degree vector of $N_j(z)$. Hence the solvability of ${}^R(V,U,W;f)_I$ is equivalent with the solvability of equation (3.13). Therefore we have proved the following theorem.

Theorem 3.1 *The problem ${}^R(V,U,W;f)_I$ is solvable if and only if*

$$\operatorname{rank} {}^RH_I^j = \operatorname{rank}\left[{}^RB_I^j \quad {}^RH_I^j\right], \quad j=1,2,\ldots,m, \tag{3.14}$$

where ${}^RB_I^j$ and ${}^RH_I^j$ are defined in (3.11) and (3.12).

In order to establish a similar theorem for the problem ${}^L(\tilde V,\tilde U,\tilde W;f)_I$ we can transform it by transposition into the problem ${}^R(\tilde V^T,\tilde U^T,\tilde W^T;f^T)_I$ and then use the above theorem. There is a simple technique which allows us to express this more directly in terms of f rather than f^T. It basically uses the fact that Toeplitz matrices are persymmetric, i.e., symmetric with respect to a SW-NE axis. We shall illustrate this technique below. The following computations will show how to derive the left versions of the matrices defined in (3.11) and (3.12) which will eventually lead to Corollary 3.2 which is the left version of Theorem 3.1 .

To emphasize the dependence on V,U,W and f, we denote temporarily ${}^RH_I^j$ and ${}^RB_I^j$ by ${}^RH_I^j(V,U,W;f)$ and ${}^RB_I^j(V,U,W;f)$ respectively. Then for the solvability of ${}^R(\tilde V^T,\tilde U^T,\tilde W^T;f^T)_I$, condition (3.14) is changed into

$$\operatorname{rank}{}^RH_I^i(\tilde V^T,\tilde U^T,\tilde W^T;f^T) = \operatorname{rank}\left[{}^RB_I^i(\tilde V^T,\tilde U^T,\tilde W^T;f^T) \quad {}^RH_I^i(\tilde V^T,\tilde U^T,\tilde W^T;f^T)\right],$$

$$i=1,2,\ldots,p.$$

Let $J_n = (\delta_{i,n+1-j})_{i,j=1}^n$, where

$$\delta_{ij} = \begin{cases} 1, & \text{for } i=j \\ 0, & \text{for } i\neq j \end{cases}$$

be the column or row reversing operator. Then the persymmetry of the Toeplitz matrices is expressed by

$$\left[J_m T_{mn}^l(g) J_n\right]^T = T_{nm}^{l+m-n}(g).$$

Hence, if we set

$$J_1 := \operatorname{diag}(J_{\tilde u_{i1}},\ldots,J_{\tilde u_{ip}}) \quad \text{and} \quad J_2 := \operatorname{diag}(J_{\tilde w_{i1}-\tilde u_{i1}},\ldots,J_{\tilde w_{im}-\tilde u_{im}}),$$

then we can obtain the left analogues of the matrices (3.11) and (3.12)

$$\left[J_1 \quad {}^RH_I^i(\tilde V^T,\tilde U^T,\tilde W^T;f^T) \ J_2\right]^T = {}^LH_I^i(\tilde V,\tilde U,\tilde W;f) = {}^LH_I^i$$

and

$$\left[J_1 \ {}^RB_I^i(\tilde V^T,\tilde U^T,\tilde W^T;f^T)\right]^T = {}^LB_I^i(\tilde V,\tilde U,\tilde W;f) = {}^LB_I^i.$$

In terms of the data, this means that they are given by

$${}^LH_I^i = \begin{bmatrix} T_{\tilde u_{i1},\tilde e_{i1}}^{\tilde w_{i1}-\tilde u_{i1}}(f_{11}) & \cdots & T_{\tilde u_{i1},\tilde e_{im}}^{\tilde w_{im}-\tilde u_{i1}}(f_{1m}) \\ \cdots & \cdots & \cdots \\ T_{\tilde u_{ip},\tilde e_{i1}}^{\tilde w_{i1}-\tilde u_{ip}}(f_{p1}) & \cdots & T_{\tilde u_{ip},\tilde e_{im}}^{\tilde w_{im}-\tilde u_{ip}}(f_{pm}) \end{bmatrix}, \tag{3.15}$$

and

$${}^LB_I^i = \left[\ T_{1,\tilde e_{i1}}^{\tilde w_{i1}}(f_{i1}) \quad \cdots \quad T_{1,\tilde e_{im}}^{\tilde w_{im}}(f_{im})\ \right], \tag{3.16}$$

while the $\tilde E$ matrix is as before defined by

$$\tilde E = (\tilde e_{ij})_{i,j=1}^{p,m} = \tilde W - \tilde V. \tag{3.17}$$

Thus the solvability condition for the left problem ${}^L(\tilde V,\tilde U,\tilde W;f)_I$ can now be formulated as in the next corollary.

Corollary 3.2 *The problem* $^L(\tilde V, \tilde U, \tilde W; f)_I$ *is solvable if and only if*

$$\operatorname{rank} {}^LH_I^i = \operatorname{rank} \begin{bmatrix} {}^LB_I^i \\ {}^LH_I^i \end{bmatrix}, \quad i = 1, 2, \ldots, p, \tag{3.18}$$

where the matrices ${}^LB_I^i$ *and* ${}^LH_I^i$ *are as defined in (3.15) and (3.16).*

In the special case where we consider row degrees for numerator and denominator and row orders, the matrix of the ssystem (3.13) becomes independent of j. Only the right hand side will be different for each column of the denominator. This is computationally an interesting situation because then computations can be done very efficently. In the following corollary, we suppose the order is constant for all entries. Then we get also a remarkable relation between a left and a right approximant.

Corollary 3.3 *Suppose the following row degrees and a uniform order of approximation for type I right MPAs are given*

$$V \in \mathbf{Z}_+^{p\times 1}, \quad U \in \mathbf{Z}_+^{m\times 1}, \quad W \in \mathbf{Z}_+,$$

and associate with these the column degrees and uniform order for corresponding left MPAs as below

$$\tilde V = (W - U)^T, \ \tilde U = (W - V)^T, \quad \tilde W = W.$$

Then the matrices ${}^RH_I^j$ *of (3.11) and* ${}^LH_I^i$ *of (3.15) are independent of* j *and* i *and they are equal to*

$${}^RH_I = \begin{bmatrix} T^{v_1}_{\tilde u_1, u_1}(f_{11}) & T^{v_1}_{\tilde u_1, u_2}(f_{12}) & \cdots & T^{v_1}_{\tilde u_1, u_m}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T^{v_p}_{\tilde u_p, u_1}(f_{p1}) & T^{v_p}_{\tilde u_p, u_2}(f_{p2}) & \cdots & T^{v_p}_{\tilde u_p, u_m}(f_{pm}) \end{bmatrix}. \tag{3.19}$$

The matrices ${}^RB_I^j$ *of (3.12) become*

$${}^RB_I^j = \begin{bmatrix} T^{v_1+1}_{\tilde u_1, 1}(f_{1j}) \\ T^{v_2+1}_{\tilde u_2, 1}(f_{2j}) \\ \vdots \\ T^{v_p+1}_{\tilde u_p, 1}(f_{pj}) \end{bmatrix}, \tag{3.20}$$

while the ${}^LB_I^i$ *of (3.16) become*

$${}^LB_I^i = \begin{bmatrix} T^w_{1,u_1}(f_{i1}) & \cdots & T^w_{1,u_m}(f_{im}) \end{bmatrix}. \tag{3.21}$$

From the general solvability conditions of 3.1 and 3.2 it follows easily that the following conclusions hold.

(i) $^R(V, U, W; f)_I$ *is solvable if and only if*

$$\operatorname{rank} {}^RH_I = \operatorname{rank}[\, {}^RB_I^j \ {}^RH_I], \quad j = 1, 2, \ldots, m; \tag{3.22}$$

(ii) $^L(\tilde V, \tilde U, \tilde W; f)_I$ *is solvable if and only if*

$$\operatorname{rank} {}^RH_I = \operatorname{rank} \begin{bmatrix} {}^LB_I^i \\ {}^RH_I \end{bmatrix} \quad i = 1, 2, \ldots, p; \tag{3.23}$$

(iii) If the matrix ${}^RH_I$ *is nonsingular, then both* $^R(V, U, W; f)_I$ *and* $^L(\tilde V, \tilde U, \tilde W; f)_I$ *are solvable.*

We note that in the case of Corollary 3.3, the existence of both sides MPA is related to the same matrix ${}^R H_I$, which does not depend on j.

So far for the existence of the approximants of type I. Now we consider the solvability of the type II problems ${}^R(V,U,W;f)_{II}$. Recall that for this type the most general case has $V, W \in \mathbf{Z}_+^{p\times m}$ and $U \in \mathbf{Z}_+^{m\times 1}$. The formulation of the results will become much more compact if we associate to the sequence $(u_i)_1^m$ three matrices which are related to the number of free parameters left at position (i,j) of the denominator when the degree conditions and the normalization are taken into account. These 3 matrices are :

$$C(U) = (c_{ij})_{i,j=1}^m, \quad T(U) = (t_{ij})_{i,j=1}^m \quad \text{and} \quad S(U) = (s_{ij})_{i,j=1}^m, \tag{3.24}$$

where

$$c_{ij} = \begin{cases} 1, & \text{for } i \geq j, \text{ or } u_i \leq u_j \text{ and } i < j \\ 0, & \text{otherwise ,} \end{cases} \tag{3.25}$$

$$t_{ij} = \min\{u_i, u_j\} - c_{ij}, \quad \text{and } s_{ij} = t_{ij} + 1. \tag{3.26}$$

The meaning of these auxiliary matrices is the following. Assume matrix $M(z)$ is $C1$ canonical with degree U. Then a closer look at the structure of a $C1$ canonical matrix shall reveal that

a. $c_{ij} = 1$ means that, in order to make $M(z)$ $C1$ canonical, there is one extra condition at the position (i,j) in determining $M(z)$, otherwise, if $c_{ij} = 0$, there is no condition.

b. s_{ij} stands for the number of unknowns in determining $M(z)$ at position (i,j).

c. t_{ij} denotes the degree of $M_{ij}(z)$ if $i \neq j$ and for $i = j$, the degree of $M_{ij}(z)$ is $t_{ij} + 1$.

Using these notations, the problem for determining the j-th column of the denominator $M(z)$ in the problem ${}^R(V,U,W;f)_{II}$ can be expressed by the following set of equations.

$${}^R H_{II}^j \, X = - \, {}^R B_{II}^j, \tag{3.27}$$

where

$${}^R H_{II}^j = \begin{bmatrix} T_{e_{1j},s_{1j}}^{v_{1j}+1}(f_{11}) & T_{e_{1j},s_{2j}}^{v_{1j}+1}(f_{12}) & \cdots & T_{e_{1j},s_{mj}}^{v_{1j}+1}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{e_{pj},s_{1j}}^{v_{pj}+1}(f_{p1}) & T_{e_{pj},s_{2j}}^{v_{pj}+1}(f_{p2}) & \cdots & T_{e_{pj},s_{mj}}^{v_{pj}+1}(f_{pm}) \end{bmatrix}, \tag{3.28}$$

and

$${}^R B_{II}^j = \begin{bmatrix} T_{e_{1j},1}^{v_{1j}-u_j+1}(f_{1j}) \\ T_{e_{2j},1}^{v_{2j}-u_j+1}(f_{2j}) \\ \vdots \\ T_{e_{pj},1}^{v_{pj}-u_j+1}(f_{pj}) \end{bmatrix}, \tag{3.29}$$

and, as for type I approximants, E is defined as $E = (e_{ij}) = W - V$. Therefore we have the following most general solvability condition for right type II approximants.

Theorem 3.4 *The problem ${}^R(V,U,W;f)_{II}$ is solvable if and only if*

$$\operatorname{rank} {}^R H_{II}^j = \operatorname{rank}[{}^R B_{II}^j \ \ {}^R H_{II}^j], \quad j = 1, 2, \ldots m, \tag{3.30}$$

where the matrices ${}^R H_{II}^j$ and ${}^R B_{II}^j$ are defined in (3.28) and (3.29).

A similar conclusion can be established for left approximants of type II, i.e., for the problem ${}^{L}(\tilde{V}, \tilde{U}, \tilde{W}; f)_{II}$. The same technique of transposing and reformulation as used to prove Corollary 3.3 for type I approximants can be here.

As for the type I MPAs, we consider the special case of row degrees for numerator and denominator and uniform order for all entries. This is done in the Corollary given below. The are some extra conditions on these row orders, given by (3.31) and (3.32). The condition (3.31) follows from the solvability equality (2.7) for type II approximants. The condition (3.32) is an extra condition which makes similar constraints for the corresponding left approximants to be satisfied. In this case, the existence conditions are completely independent of the row or column index.

Corollary 3.5 *Consider row degrees for numerator and denominator and a uniform order for right type II MPAs, i.e.,*

$$V \in \mathbf{Z}_+^{p\times 1},\ U \in \mathbf{Z}_+^{m\times 1},\quad W \in \mathbf{Z}_+,$$

and suppose they satisfy the following conditions

$$u_1 \geq u_2 \geq \cdots \geq u_m, \quad |u_1 - u_m| \leq 1, \tag{3.31}$$

and

$$v_1 \leq v_2 \leq \cdots \leq v_p, \quad |v_p - v_1| \leq 1. \tag{3.32}$$

For the corresponding left approximants we consider the column degrees and uniform order, related to the previous ones by

$$\tilde{V} = (W - U)^T,\ \tilde{U} = (W - V)^T,\quad \tilde{W} = W.$$

The matrices ${}^{R}H_{II}^{j}$ become independent of j and are now given by (we indicate the dependency on V, U and W explicitly)

$${}^{R}H_{II}(V,U,W) = \begin{bmatrix} T_{\tilde{u}_1,u_1}^{v_1+1}(f_{11}) & T_{\tilde{u}_1,u_2}^{v_1+1}(f_{12}) & \cdots & T_{\tilde{u}_1,u_m}^{v_1+1}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{\tilde{u}_p,u_1}^{v_p+1}(f_{p1}) & T_{\tilde{u}_p,u_2}^{v_p+1}(f_{p2}) & \cdots & T_{\tilde{u}_p,u_m}^{v_p+1}(f_{pm}) \end{bmatrix}. \tag{3.33}$$

The existence condition for the right MPAs given in Theorem 3.4 and of its left analogue can now be formulated as follows :

(i) ${}^{R}(V,U,W;f)_{II}$ *is solvable if and only if*

$$\operatorname{rank} {}^{R}H_{II}(V,U,W) = \operatorname{rank} {}^{R}H_{II}(V,U+1,W), \tag{3.34}$$

(ii) ${}^{L}(\tilde{V},\tilde{U},\tilde{W};f)_{II}$ *is solvable if and only if*

$$\operatorname{rank} {}^{R}H_{II}(V,U,W) = \operatorname{rank} {}^{R}H_{II}(V-1,U,W), \tag{3.35}$$

(iii) If the matrix ${}^{R}H_{II}(V,U,W)$ *is nonsingular, then both* ${}^{R}(V,U,W;f)_{II}$ *and* ${}^{L}(\tilde{V},\tilde{U},\tilde{W};f)_{II}$ *are solvable.*

Proof. (i) One can easily check that under condition (3.31), $s_{ij} = u_i$. This implies, as we mentioned already, that the matrices ${}^{R}H_{II}^{j}$ of (3.28) become independent of j and are all equal to the matrix ${}^{R}H_{II}(V,U,W)$ as defined in (3.33). On the other hand, (3.30) then says that for any j, the matrix ${}^{R}B_{II}^{j}$ defined as in (3.29) should not add to the rank of ${}^{R}H_{II} = {}^{R}H_{II}(V,U,W)$. This implies that we should have

$$\begin{aligned} \operatorname{rank} {}^{R}H_{II}(V,U,W) &= \operatorname{rank}[{}^{R}B_{II}^{1} \cdots {}^{R}B_{II}^{m}\ {}^{R}H_{II}^{j}(V,U,W)] \\ &= \operatorname{rank} {}^{R}H_{II}(V,U+1,W), \end{aligned}$$

as you can see after rearrangement of the columns. This proves (i).

(ii) The second part of the conclusion can be proved by an approach very similar to the one used in Corallary 3.2. We shall not repeat it here.

(iii) The third conclusion follows from (i) and (ii). □

Besides the results given in this section, many other similar results can be obtained by considering special cases. We leave it to the devotion of the reader to formulate some other results which suit his special interest for a certain type of MPAs. Setting up the equations explicitly will help him to give explicit determinant expressions like we shall do in the next section and it will certainly help him to design his favorite computational scheme. If not, these equations will at least tell him how to compute the MPAs in principle.

4 Determinant Expressions for MPAs

In all books on scalar Padé approximation one can find explicit formulas, so called determinant formulas, that give explicit expressions for numarator and denominator as the ratio of two determinants that depend on the given power series. In this section we want to give determinant expressions for the numerator and denominator of the elements in ${}^{R}[V,U,W;f]_{I}$ and ${}^{R}[V,U,W;f]_{II}$ i.e., of right MPAs of type I and type II. We do this only in the most general setting. The reader can adapt these expressions to his favored special case. Because these results follow immediately from the equations which were derived in the previous section, we shall reuse the notations for the matrices ${}^{R}H_{I/II}$ and ${}^{R}B_{I/II}$ introduced already in (3.11), (3.12), (3.28) and (3.29). Since only right approximants are considered in this section, we shall leave out the R-superscript. Besides these notations, we need a way of denoting truncated series as we shall presently introduce. For a given power series

$$g(z)=\sum_{i=0}^{\infty}a_iz^i \quad a_i\in\mathbf{C},$$

let

$$g^{(k)}(z)=\sum_{i=0}^{k}a_iz^i, \quad a_i\in\mathbf{C},$$

and

$$\bar{g}^{(k)}(z)=\sum_{i=k+1}^{\infty}a_iz^i, \quad a_i\in\mathbf{C}.$$

If $g(z)=(g_{ij}(z))_{i,j=1}^{s,t}$ and $V\in\mathbf{Z}_{+}^{s\times t}$, then we denote $(g_{ij}^{(v_{ij})}(z))_{i,j=1}^{s,t}$ and $(\bar{g}_{ij}^{(v_{ij})}(z))_{i,j=1}^{s,t}$ by $g^{(V)}(z)$ and $\bar{g}^{(V)}(z)$ respectively. The zero vector of k entries is denoted as :

$$\theta_k=[0,0,\ldots,0]\in\mathbf{C}^{1\times k}.$$

We are now ready to give the determinant expressions for the right type I MPAs.

Theorem 4.1 *Let $N(z)M(z)^{-1}\in{}^{R}[V,U,W;f]_I$ be a type I right MPA for general V,U and W. Then, if the matrix $H_I^j={}^{R}H_I^j$ (see (3.11)) is nonsingular and $B_I^j={}^{R}B_I^j$ is defined as in (3.16), we can express the (i,j)-th element of the denominator as*

$$M_{ij}(z)=\frac{1}{\det H_I^j}\det\begin{bmatrix}\delta_{ij} & \lambda_I(i,j)\\ B_I^j & H_I^j\end{bmatrix}, \tag{4.36}$$

while the (i,j)-th element of the numerator is expressed as

$$N_{ij}(z)=\frac{1}{\det H_I^j}\det\begin{bmatrix}f_{ij}^{(v_{ij})}(z) & \omega_I(i,j)\\ B_I^j & H_I^j\end{bmatrix}, \tag{4.37}$$

where

$$\lambda_I(i,j) = [\theta_{u_{1j}}, \ldots, \ \theta_{u_{i-1,j}}, \ z, \ldots, z^{u_{ij}}, \ \theta_{u_{i+1,j}} \ , \ldots, \theta_{u_{m,j}}],$$

$$\omega_I(i,j) = [\delta_I(i,j,1), \ldots, \delta_I(i,j,m)],$$

and

$$\delta_I(i,j,l) = [z f_{il}^{(v_{ij}-1)}(z), \ldots, z^{u_{lj}} f_{il}^{(v_{ij}-u_{lj})}(z)].$$

Proof. The elements $M_{ij}(z)$ and $N_{ij}(z)$ as expressed by the determinants above have the correct degrees u_{ij} and v_{ij} respectively. Also the normalization condition $M(0) = I$ is easily checked. We shall prove that the (i,j)-th element $R_{ij}(z)$ of $R(z) = f(z)M(z) - N(z)$ with the entries of $M(z)$ and $N(z)$ as defined in (4.36) and (4.37) has the correct order w_{ij}. Then we have proved the theorem. As a side result we also have a determinant expression for the residual entries $R_{ij}(z)$.

$$\begin{aligned} R_{ij}(z) &= \sum_{k=1}^{m} f_{ik}(z) M_{kj}(z) - N_{ij}(z) \\ &= \frac{1}{\det H_I^j} \det \begin{bmatrix} f_{ij}(z) - f_{ij}^{(v_{ij})}(z) & \sum_{k=1}^{m} f_{ik}(z)\lambda_I(k,j) - \omega_I(i,j) \\ B_I^j & H_I^j \end{bmatrix} \\ &= \frac{1}{\det H_I^j} \det \begin{bmatrix} \bar{f}_{ij}^{(v_{ij})}(z) & z\bar{f}_{i1}^{(v_{ij}-1)} \ldots z^{u_{1j}} \bar{f}_{i1}^{(v_{ij}-u_{1j})}(z) \ z\bar{f}_{i2}^{(v_{ij}-1)} \ldots \\ B_I^j & H_I^j \end{bmatrix} \end{aligned}$$

We can eliminate a number of terms from the first row by subtracting multiples of the next rows in the above determinant:

$$\begin{aligned} R_{ij}(z) &= \frac{1}{\det H_I^j} \det \begin{bmatrix} \bar{f}_{ij}^{(w_{ij})}(z) & z\bar{f}_{i1}^{(w_{ij}-1)} \ldots z^{u_{1j}} \bar{f}_{i1}^{(w_{ij}-u_{1j})}(z) \ z\bar{f}_{i2}^{(w_{ij}-1)} \ldots \\ B_I^j & H_I^j \end{bmatrix} \\ &= O(z^{w_{ij}+1}). \end{aligned}$$

Now the theorem is proved. □

For the second type MPA, similar formulas can be established.

Theorem 4.2 *Let $N(z)M(z)^{-1} \in {}^R[V,U,W;f]_{II}$ be a right type II MPA. Then, if the matrix $H_{II}^j = {}^R H_{II}^j$ as defined in (3.28) is nonsingular and if $B_{II}^j = {}^R B_{II}^j$ is as defined in (3.29), then we can express the elements of the denominator by the following determinant formula*

$$M_{ij}(z) = \frac{1}{\det H_{II}^j} \det \begin{bmatrix} \delta_{ij} z^{s_{jj}} & \lambda_{II}(i,j) \\ B_{II}^j & H_{II}^j \end{bmatrix}, \tag{4.38}$$

and the numerator coefficients are given by

$$N_{ij}(z) = \frac{1}{\det H_{II}^j} \det \begin{bmatrix} z^{s_{jj}} f_{ij}^{(v_{ij}-s_{jj})}(z) & \omega_{II}(i,j) \\ B_{II}^j & H_{II}^j \end{bmatrix}, \tag{4.39}$$

where

$$\lambda_{II}(i,j) = [\theta_{s_{1j}}, \ldots, \theta_{s_{i-1,j}}, 1, \ z \ , \ldots, z^{t_{ij}}, \ \theta_{s_{i+1,j}}, \ldots, \theta_{s_{m,j}}],$$

$$\omega_{II}(i,j) = [\delta_{II}(i,j,1), \ldots, \delta_{II}(i,j,m)],$$

and

$$\delta_{II}(i,j,l) = [f_{il}^{(v_{ij})}(z), \ldots, z^{t_{lj}} f_{il}^{(v_{ij}-t_{lj})}(z)].$$

Recall that the numbers s_{ij} and t_{ij} are as defined in (3.24)-(3.26).

Proof. The proof is very similar to the previous one. First note that from (4.38) we can see that the degree of $M_{jj}(z)$ is given by $\partial M_{jj} = s_{jj} = u_j$, and that the coefficient of z^{u_j} is one. For $i \neq j$ the degrees of the M_{ij} satisfy

$$\partial M_{ij} \leq t_{ij}.$$

It follows by the construction of $T(U)$ and $S(U)$ (see section 3) that $M(z)$ is $C1$ canonical. Also the degrees for the $N_{ij}(z)$ entries is bounded by v_{ij}. Thus only the order of the residual entries remains to be checked. Calculate the (i,j)-th element $R_{ij}(z)$ of $R(z) = f(z)M(z) - N(z)$

$$\begin{aligned}
R_{ij}(z) &= \sum_{k=1}^{m} f_{ik}(z)M_{kj}(z) - N_{ij}(z) \\
&= \frac{1}{\det H_{II}^{j}} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(v_{ij}-s_{jj})}(z) & \sum_{k=1}^{m} f_{ik}(z)\lambda_{II}(k,j) - \omega_{II}(i,j) \\ B_{II}^{j} & H_{II}^{j} \end{bmatrix} \\
&= \frac{1}{\det H_{II}^{j}} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(v_{ij}-s_{jj})}(z) & \bar{f}_{i1}^{(v_{ij})} \cdots z^{t_{1j}} \bar{f}_{i1}^{(v_{ij}-t_{1j})}(z) \;\; \bar{f}_{i2}^{(v_{ij})} \cdots \\ B_{II}^{j} & H_{II}^{j} \end{bmatrix}
\end{aligned}$$

Again we find that by row combinations we can eliminate some of the terms in the first row to find

$$\begin{aligned}
R_{ij}(z) &= \frac{1}{\det H_{II}^{j}} \det \begin{bmatrix} z^{s_{jj}} \bar{f}_{ij}^{(w_{ij}-s_{jj})}(z) & \bar{f}_{i1}^{(w_{ij})} \cdots z^{t_{1j}} \bar{f}_{i1}^{(w_{ij}-t_{1j})}(z) \;\; \bar{f}_{i2}^{(w_{ij})} \cdots \\ B_{II}^{j} & H_{II}^{j} \end{bmatrix} \\
&= O(z^{w_{ij}+1}).
\end{aligned}$$

Now the theorem is proved. □

We leave it to the reader to find similar formulas for the left MPAs and to investigate special cases.

5 Relations Among the Different MPAs

As we already have mentioned in section 2, the type I and type II approximants are basically different concepts and were proposed on a different basis. Thus there is little hope, except of course in the scalar case, that these two definitions are stronly related. The relation between matrix minimal Padé approximants and type II approximants is somewhat more to be expected since they have the same normalization. The problem we consider in this section is to find some relations among the different definitions of MPA. We try to answer the questions : do the two types of MPAs define the same approximant? and is the solution of problem MmPA a solution of problem ${}^{R}(V,U,W;f)_{I}$ or ${}^{R}(V,U,W;f)_{II}$? We shall show that, in general, the different definitions lead to different approximants. Only in some special cases, they determine the same one. We shall start with relations that can be given between type I and type II approximants. The relation with MmPAs is given lateron.

When we were discussing the possible normalizations, we have already warned for the fact that if we renormalize by multiplication with a right unimodular factor, we mix up the degree structure completely. Since changing a type I MPA into a type II MPA or conversely requires a renormalization, it is obvious that the best we can expect is some correspondence if we consider row degrees and orders. If we then allow only a constant factor from the right to obtain this renormalization, we can be sure that the degree structure is kept. Therefore, the next Theorem gives about the most general result that can be expected in this direction.

Theorem 5.1 *Let $V, W \in \mathbf{Z}_{+}^{p\times 1}$ and $U \in \mathbf{Z}_{+}^{m\times 1}$ denote as usual the row degrees and orders of right type I or type II MPA problems. Then the following inclusion properties can be proved.*

(i) If solvability equality (2.5) for the type II approximants holds and

$$|u_i - u_j| \leq 1, \quad \forall (i,j), \tag{5.40}$$

$$u_1 \geq u_2 \geq \cdots \geq u_m, \tag{5.41}$$

then the type II approximants with $M(0)$ nonsingular can be renormalized as type I approximants for the same function. More precisely

$$\{NM^{-1} \in {}^{R}[V,U,W;f]_{II} : M(0) \text{ is nonsingular}\} \subset {}^{R}[V,U,W;f]_{I}.$$

(ii) If solvability equality (2.4) for type I approximants holds and if the denominator of the type I approximant $NM^{-1} \in {}^{R}[V,U,W;f]_I$ can be made $C1$ canonical of degree U by a right constant factor, then it will by this renormalization become a type II approximant. There is however a slight flaw for the numerator degrees. In precise terms we have

$$\{NM^{-1} \in {}^{R}[V,U,W]_I^f : \exists\, Q \in \mathbf{C}^{m\times m} \text{ s.t. } MQ \text{ is } C1 \text{ canonical with degree } U\}$$

$$\subset {}^{R}[V',U,W]_{II}^f,$$

where (V',U,W) satisfies (2.5) and $V' \geq V$.

Proof. (i). Under the conditions (5.40) and (5.41), it follows from Lemma 2.1 that the solvability equality (2.5) becomes (2.4). On the other hand, note that we consider row degrees for N and M and row orders for $fM - N$, and that these do not change when the matrices are multiplied from the right with $M(0)^{-1}$. Therefore conclusion (i) holds.

(ii). For the same V, U and W, the right hand side of (2.4) is larger than or equal to the right side of (2.5). Hence if (2.4) holds, then (2.5) holds for (V',U,W) with $V' \geq V$. Therefore (ii) is true. □

By inverting the order of the coefficients, we can also exchange the role of the leading coefficients the constant terms of the denominator. This may be a way on transforming a type I into a type II approximant. If we take for simplicity a uniform degree $U = u \in \mathbf{Z}_+$ for the denominator, then such a reversion is obtained by $M(z^{-1})z^u$. If this transformation is applied to the approximant $N(z)M(z)^{-1}$ to give the approximant $[N(z^{-1})z^u][M(z^{-1})z^u]^{-1}$, then it becomes an approximant of the series $f(z^{-1})z^u$, which is not a power series anymore. Thus we have to make a correction for this problem. This is possible if we make use of the fact that the approximant depends only on $f^{(W)}(z)$, so that we may as well suppose that $\bar{f}^{(W)}(z) = 0$. If $W \in \mathbf{Z}_+^{p\times 1}$ denotes row orders, then we can make use of the transformation $\mathrm{diag}[z^{w_1},\ldots,z^{w_p}]f(z^{-1})$. The next theorem uses these ideas and shows that in this way a type I approximant can be transformed into a type II approximant and conversely.

Theorem 5.2 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U = u \in \mathbf{Z}_+$ be given such that solvability equality (2.4) for type I approximants holds. (Note that in this case (2.4) is the same as (2.5) so that also the solvability equality for type II approximants holds.) For the given $f(z)$, define*

$$g(z) = (g_{ij})_{i,j=1}^{p,m},$$

where

$$g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1}),$$

which means that the order of the coefficients of $f^{(W)}(z)$ is reversed. Suppose furthermore that there exists a type II solution for $f(z)$, i.e.,

$${}^{R}[V,U,W;f]_{II} \neq \emptyset.$$

Then we have the following conversion possibilities between type I and type II MPAs :

(i) Any type II MPA of f

$$N_2(z)M_2(z)^{-1} \in {}^R[V,U,W;f]_{II},$$

can be transformed into a type I approximant for g. More precisely : if we define

$$M_1(z) = M_2(z^{-1})z^u,$$

and

$$N_1(z) = g(z)M_2(z^{-1})z^u - \mathrm{diag}[z^{w_1},\ldots,z^{w_p}]N_2(z^{-1})z^u,$$

then

$$N_1(z)M_1(z)^{-1} \in {}^R[U-1,U,W+U-V-1;g]_I.$$

(ii) Any type I MPA for g

$$N_1(z)M_1(z)^{-1} \in {}^R[U-1,U,W+U-V-1;g]_I,$$

can be transformed into a type II solution for f, i.e., if we set

$$M_2(z) = M_1(z^{-1})z^u,$$

and

$$N_2(z) = f^{(W)}(z)M_1(z^{-1})z^u - \mathrm{diag}[z^{w_1},\ldots,z^{w_p}]N_1(z^{-1})z^u,$$

then

$$N_2(z)M_2(z)^{-1} \in {}^R[V,U,W;f]_{II}.$$

Proof. (i). Since the problem ${}^R(V,U,W;f)_{II}$ is independent of $\bar{f}_{ij}^{(w_i)}(z)$, we may consider it as being zero and replace $f_{ij}(z)$ by $f_{ij}^{(w_i)}(z)$. Let $N_2(z)M_2(z)^{-1} \in {}^R[V,U,W;f]_{II}$. Then

$$f(z)M_2(z) - N_2(z) =: R(z) \in \mathbf{E}_W^{p\times m} \cap \mathbf{H}_{U+W}^{p\times m}. \tag{5.42}$$

Now set $z = t^{-1}$ and multiplying by $\mathrm{diag}[t^{w_1},\ldots,t^{w_p}]$ on the left and by t^u on the right, then we get

$$g(t)M_2(t^{-1})t^u - \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u = N_1(t), \tag{5.43}$$

where

$$\begin{aligned} N_1(t) &= \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]R(t^{-1})t^u \\ &= g(t)M_2(t^{-1})t^u - \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u. \end{aligned} \tag{5.44}$$

Let

$$M_1(t) = M_2(t^{-1})t^u.$$

Then (5.43) can be written as

$$g(t)M_1(t) - N_1(t) = \mathrm{diag}[t^{w_1},\ldots,t^{w_p}]N_2(t^{-1})t^u. \tag{5.45}$$

Since $M_2(z)$ is $C1$ canonical it follows that $M_1(0) = I$. From (5.42) and (5.45) we also find that

$$N_1(t) \in \mathbf{H}_{U-1}^{p\times m}$$

and

$$g(t)M_1(t) - N_1(t) \in \mathbf{E}_{W+U-V-1}^{p\times m}.$$

Finally we can note that because of (2.4)

$$\begin{aligned}\sum_{i=1}^{p}(w_i + u_i - v_i - 1) &= \sum_{i=1}^{p} w_i + p(u-1) - \sum_{i=1}^{p} v_i \\ &= mu + p(u-1).\end{aligned}$$

i.e., the solvablity equality holds for the problem ${}^R(U-1, U, W+U-V-1; g)_I$. Hence

$$N_1(t)M_1(t)^{-1} \in {}^R[U-1, U, W+U-V-1; g]_I.$$

(ii). Conclusion (ii) can be proved by a similar approach. □

Now we turn our interests to finding some relations between the type I and type II definitions of MPAs given in this paper and the minimal Padé approximants in the matrix case (MmPA). The MmPA problem has been studied extensively in [7]. The (minimal) degrees for the denominator columns (and rows) of the solutions are known to be so called Kronecker indices. The row degrees of the corresponding numarator are related to dual Kronecker indices. Since these numbers are very common in the system literature, but probably less popular in the Padé literature ([15],[1],[17]), it is worthwhile to repeat their definition here.

Definition 5.1 (Kronecker indices) *Given a sequence $a = \{a_0, a_1, \ldots, a_\mu\}$ of elements from $\mathbf{C}^{p\times m}$. Consider for $j = 0, 1, \ldots, \mu$ the matrices*

$$A^{(j)} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{\mu-j} \\ a_1 & a_2 & \cdots & a_{\mu-j+1} \\ \cdots & \cdots & \cdots & \cdots \\ a_j & a_{j+1} & \cdots & a_\mu \end{bmatrix}$$

The dual Kronecker indices of the sequence a are defined for $i = 1, \ldots, p$ by

$$\tilde{u}_i = \begin{cases} \min\{j : \textit{row } jp+i \textit{ is lin. dep. on the previous rows of } A^{(j)}\} \\ \mu+1 \qquad \textit{if there is no linear dependency for any } j \end{cases}$$

The m Kronecker indices are similarly defined expressing linear dependencies for the columns of the matrices

$$\tilde{A}^{(j)} = \begin{bmatrix} a_0 & a_1 & \cdots & a_j \\ a_1 & a_2 & \cdots & a_{j+1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{\mu-j} & a_{\mu-j+1} & \cdots & a_\mu \end{bmatrix}.$$

The Kronecker and dual Kronecker indices for the empty sequence are supposed to be zero.

What is it than one can expect as a theorem? In accordance with the definition of MmPA and the minimality property of the Kronecker indices, we may hope to prove that a MmPA is a MPA if we choose for denominator column (and row) degrees the Kronecker indices, the order to be uniform and the row degrees of the numerator as prescribed by the definition of MmPA. The only thing that needs to be checked then is that $\det M(0) \neq 0$ for the type I approximants and the solvability equality (2.5) for the type II approximants. This is formulated in the next theorem.

Theorem 5.3 *Let the power series $f(z) = \sum_{k=0}^{\infty} c_k z^k \in \mathbf{C}^{p\times m}[[z]]$ be given, as well as the integers $\omega \in \mathbf{Z}_+$ and $\delta \in [-\omega, \omega]$. Let $U = (u_i)_{i=1}^m \in \mathbf{Z}^{m\times 1}$ and $\tilde{U} = (\tilde{u}_i)_{i=1}^p \in \mathbf{Z}^{p\times 1}$ be the Kronecker and dual Kronecker indices respectively associated with the sequence $\{c_\omega, c_{\omega-1}, \ldots, c_{-\delta+1}\}$. Finally suppose that NM^{-1} denotes a solution of problem MmPA for the given ω and δ. Then the following is true :*

(i) If

$$V = \omega - \tilde{U} \quad \text{and} \quad W = \omega, \tag{5.46}$$

then the solvability equality (2.4) holds for V, U and W. Moreover NM^{-1} is a type I MPA, i.e.,

$$NM^{-1} \in {}^R[V, U, W; f]_I,$$

provided $M(0)$ is nonsingular.

(ii) For any $V, W \in \mathbf{Z}_+^{p\times 1}$, such that

$$\omega - \tilde{U} \le V \le W \le \omega, \tag{5.47}$$

suppose that the solvability equality (2.5) holds for V, U, W. In that case the MmPA solution NM^{-1} will be a type II MPA i.e.,

$$NM^{-1} \in {}^R[V, U, W; f]_{II}.$$

Proof. (i). It is a very well known property that the sum of the Kronecker indices is equal to the sum of the dual Kronecker indices. Thus that

$$\sum_{i=1}^{p} \tilde{u}_i = \sum_{i=1}^{m} u_i.$$

This equality implies that

$$\sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i = p\omega - \sum_{i=1}^{p} \tilde{u}_i + \sum_{i=1}^{m} u_i = p\omega.$$

i.e., (2.4) holds, it follows from [7] that the degree of the i-th row of N is less than or equal to $\omega - \tilde{u}_i = v_i$. Thus $(N, M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m}$ and comonic renormalization by a multiplication from the right with $M(0)^{-1}$ will not disturb this degree structure. Hence (i) is valid.

(ii). We have mentioned before that, for the same V and U, the right hand side of the type I solvability equality (2.4) is not less than the right hand side of type II solvability equality (2.5). From (i) we know that, for V and W satisfying (5.46) the equality (2.4) holds for V, U and W. Therefore in order to make (2.5) to become true, we take V and W as in (5.47). This is always possible. Because M is $C1$ canonical, it trivially follows that $NM^{-1} \in {}^R[V, U, W; f]_{II}$. □

Because the type II MPAs and the solutions of the problem MmPA both have a canonical normalization, one can expect that these two problems are much more related than as given in the previous theorem. The difficulty is that the solvability equality (2.5) has to be satisfied, since this is part of the definition of type II MPAs, while such a condition does not appear in the MmPA problem. On the other hand, the MmPA problem always has a solution, while for the type II MPA problem, for certain V, U and W, the problem may not have a solution. Thus in order to get an equivalence between the two problems, we have to set up certain V, U, W which correspond to Kronecker indices and at the same time satisfy (2.5). The existence of the type II approximants is guaranteed by requiring the normality of f. In the next theorem we shall show that the problems MmPA and type II MPA are equivalent under this normality condition of the function.

In what follows $\lfloor a \rfloor$ will denote the largest integer not exceeding a and $\lceil a \rceil$ the smallest integer larger than or equal to a.

We start with the definition of a normal power series.

Definition 5.2 (normal) *Let* $f(z) = \sum_{k=0}^{\infty} c_k z^k$ *be given. We say that* $f(z)$ *is normal if all the matrices* H_k^ω *are nonsingular for any nonnegative integer* $\omega \in \mathbf{Z}_+$ *and* $k > 0$*, satisfying*

$$\lceil k/m \rceil \leq \omega, \quad \lceil k/p \rceil \leq \omega,$$

where H_k^ω *is a* $k \times k$ *principal minor of the block Hankel matrix with entries* $c_\omega, c_{\omega-1}, c_{\omega-2}, \ldots$*, i.e.,* H_k^ω *is the matrix defined by*

$$H_k^\omega = \begin{bmatrix} c_\omega & \cdots & c_{\omega-q_1+1} & c_{\omega-q_1}(p,s) \\ \cdots & \cdots & \cdots & \cdots \\ c_{\omega-q_2+1} & \cdots & c_{\omega-q_1-q_2+2} & c_{\omega-q_1-q_2+1}(p,s) \\ c_{\omega-q_2}(t,m) & \cdots & c_{\omega-q_1-q_2+1}(t,m) & c_{\omega-q_1-q_2}(t,s) \end{bmatrix},$$

with

$$k = q_1 m + s, \quad s < m,$$

$$k = q_2 p + t, \quad t < p,$$

and $c_l(i,j)$ *denotes a matrix formed from the first* i *rows and the first* j *columns of* c_l*.*

Note that in the scalar case for $p = m = 1$ the above definition reduces to the classical notion of normality. It is a normality of the complete Padé table. For the next theorem it was sufficient to have a local normality which depends on ω and δ. This would however complicate the notation and formulation considerably and we chose not to do so.

We know from [7], and we used this in the previous theorem already, that the degrees of numerator and denominator of an MmPA with parameters (ω, δ) are related to the Kronecker and dual Kronecker indices of the sequence $c_\omega, c_{\omega-1}, \ldots, c_{-\delta+1}$. To find these indices, one has to check ranks of matrices like H_k^ω and since these are all nonsingular, these indices can be found easily. It is basically a matter of counting the blocks which is done by the parameters μ and $\tilde{\mu}$ in the next theorem.

Theorem 5.4 *For the given* (ω, δ) *as in the MmPA problem, i.e.,* ω *a nonnegative integer and* δ *an integer from* $[-\omega, \omega]$*. let*

$$\lambda = \max\{0, -\delta + 1\},$$

$$\mu = \lfloor \frac{p(\omega - \lambda + 1) - m}{p + m} \rfloor, \tag{5.48}$$

$$\tilde{\mu} = \lfloor \frac{m(\omega - \lambda + 1) - p}{p + m} \rfloor, \tag{5.49}$$

$$s = \max\{0, p(\omega - \lambda + 1) - (\mu + 1)(p + m)\},$$

and

$$t = \max\{0, m(\omega - \lambda + 1) - (\tilde{\mu} + 1)(p + m)\}.$$

Assume $f(z)$ *is normal. Then the problem MmPA is equivalent to the problem* ${}^R(V, U, W)_{II}^f$*, where*

$$V \in \mathbf{Z}_+^{p \times 1}, \quad U \in \mathbf{Z}_+^{m \times 1}, \quad \text{and} \quad W \in \mathbf{Z}_+$$

are defined by

$$v_i = \omega - \tilde{u}_i, \quad i = 1, 2, \ldots, p,$$

$$u_i = \begin{cases} \mu + 2, & i = 1, 2, \ldots, s, \\ \mu + 1, & i = s + 1, \ldots, m, \end{cases} \tag{5.50}$$

$$\tilde{u}_i = \begin{cases} \tilde{\mu}+2, & i = 1,2,\ldots,t, \\ \tilde{\mu}+1, & i = t+1,\ldots,p, \end{cases} \tag{5.51}$$

and

$$W = \omega.$$

Proof. If the definitions of u_i and $\tilde{u}_i$ have to make sense, we should first show that $s < m$ and $t < p$. Since it follows from (5.48) that there exists an integer r such that

$$p(\omega - \lambda + 1) - m = \mu(p+m) + r, \quad r < p+m,$$

we also may write

$$p(\omega - \lambda + 1) - (\mu+1)(p+m) = r - p < m.$$

Hence $s < m$. Similarly, $t < p$ can be shown.

It is easy to see that under the normality condition of f, $(u_i)_{i=1}^m$ and $(\tilde{u}_i)_{i=1}^p$ defined as in (5.50) and (5.51) are Kronecker indices and dual Kronecker indices of the sequence $c_\omega, \ldots, c_{-\delta+1}$. Hence, it follows as in the proof of Theorem 5.3 that solvability equality (2.4) holds for V, U and W as defined above. It follows then from Lemma 2.1 that (2.5) also holds. Therefore the solution of problem MmPA is also a type II MPA for these V, U, W, i.e., it also belongs to ${}^R[V, U, W; f]_{II}$.

On the other hand, by reordering the rows and columns, the matrix ${}^R H_{II}^j$ of (3.28), which is the coefficient matrix of the linear system of equations determining the denominator of the type II MPAs ${}^R[V, U, W; f]_{II}$, can be transformed into H_k^ω, where $k = \sum_{i=1}^m u_i$. Hence, under the assumption of normality, the solution of the type II MPA problem ${}^R(V, U, W; f)_{II}$ always exists uniquely. Consequently ${}^R[V, U, W]_{II}^f$ is also the (unique) solution of problem MmPA.

This completes the proof of the equivalence. □

Note that the normality condition is rather restrictive. It was already for the scalar case and it is even more so for the matrix case.

6 Properties of MPAs

For the scalar case one can find in most textbooks some invariance properties under certain linear fractional transforms. For the square matrix case with uniform degrees and orders, such type of properties can also be found in volume 2 of [2] (like e.g., Property 6.2). In this section we derive some of these properties of MPAs in the general case. Since only right approximants appear in this section we shall drop the indication ^{R} from the notation. A subscript ${}_{I,II}$ shall mean one of both ${}_I$ or ${}_{II}$. Note that in the properties to follow we had to be very careful, not only for the solvability equalities to hold, but also to preserve the degree structure or at least check how it changes by the transformations. That is why the properties are only proved for certain degree/order structures and not for the most general case. The first two properties are very simple to prove and are only mentioned for completeness without including the proof.

Property 6.1 *Let $A \in \mathbf{C}^{p\times p}$ be nonsingular. If $V, W \in \mathbf{Z}_+^{1\times m}$ and $U \in \mathbf{Z}_+^{m\times m}$, then*

$$[V, U, W; Af]_{I,II} = A\ [V, U, W; f]_{I,II}.$$

Property 6.2 *Let $A \in \mathbf{C}^{p\times p}$ and $B \in \mathbf{C}^{m\times m}$ be nonsingular. If V, U and W are all in $\mathbf{Z}_+$, then*

$$[V, U, W; AfB]_{I,II} = A\ [V, U, W; f]_{I,II}\ B.$$

In the following property we consider a linear fractional transformation of the variable z for MPAs of type I with column degrees and orders.

Property 6.3 *Consider the fractional transform of the variable* y

$$z = \frac{\alpha y}{1+\beta y}, \quad \alpha \neq 0, \quad \textit{and set} \quad g(y) = f(z).$$

Then, if the column degrees of numerator and denominator are equal, i.e., $V = U \in \mathbf{Z}_+^{1\times m}$ *and the column order is denoted by* $W \in \mathbf{Z}_+^{1\times m}$*, the following equality holds for the type I MPAs*

$$[V, U, W; g]_I(y) = [V, U, W; f]_I(z). \tag{6.52}$$

Proof. Let $N(z)M(z)^{-1} \in [V, U, W; f]_I$ be a type I MPA of f. Then by definition the folowing two relations hold

$$f(\frac{\alpha y}{1+\beta y})M(\frac{\alpha y}{1+\beta y}) - N(\frac{\alpha y}{1+\beta y}) = O\left((\frac{\alpha y}{1+\beta y})^{W+1}\right),$$

$$M(0) = I.$$

Multiplying the first equality on the right by $D(U) = \text{diag}[(1+\beta y)^{u_1}, \ldots, (1+\beta y)^{u_m}]$, we have

$$g(y)q(y) - p(y) = O(y^{W+1}),$$

$$q(0) = I,$$

where

$$q(y) = M(\frac{\alpha y}{1+\beta y})D(U), \quad p(y) = N(\frac{\alpha y}{1+\beta y})D(U).$$

Hence

$$p(y)q(y)^{-1} = N(z)M(z)^{-1} \in [V, U, W; g]_I$$

is a type I MPA for g. Thus we have the inclusion

$$[V, U, W; f]_I(z) \subset [V, U, W; g]_I(y).$$

Similarly, we can prove the converse inclusion

$$[V, U, W; g]_I(y) \subset [V, U, W; f]_I(z),$$

so that equality (6.52) holds. □

The following property says that if a polynomial of a sufficiently low degree is added to f, then the type I or II MPA also gets this polynomial added. This property holds for the most general degree and order situation.

Property 6.4 *Let* $V, W, E \in \mathbf{Z}_+^{p\times m}$ *and* $U \in \mathbf{Z}_+^{m\times m}$*. If*

$$\max_{1\le k\le m} \{e_{ik} + u_{kj}\} \le v_{ij}, \quad i = 1, \ldots, p; \quad j = 1, \ldots, m, \tag{6.53}$$

then for any polynomial matrix $B(z) \in \mathbf{H}_E^{p\times m}$*, we have*

$$[V, U, W; f]_{I,II} + B(z) = [V, U, W; f + B]_{I,II}. \tag{6.54}$$

Proof. Since the subscript $_{I,II}$ is irelevant in the proof, we shall drop it from the notation. Let $N(z)M(z)^{-1} \in [V, U, W; f]$ be some MPA for f. Then the order of approximation is given by W, i.e.,

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}.$$

Hence also

$$[f(z)+B(z)]M(z)-[N(z)+B(z)M(z)] \in \mathbf{E}_W^{p\times m}.$$

The condition (6.53) on the degree E of $B(z)$ expresses that $N(z)+B(z)M(z) \in \mathbf{H}_V^{p\times m}$ is a polynomial matrix of degree at most V. Therefore

$$[N(z)+B(z)M(z)]M(z)^{-1} = N(z)M(z)^{-1}+B(z) \in [V,U,W;f+B]$$

is a MPA for $f+B$. In other words, we have proved the inclusion

$$[V,U,W;f]+B(z) \subset [V,U,W;f+B].$$

The opposite direction can be proved similarly. □

The next property shows how MPAs for successive parts of the expansion of f can be combined to give a MPA for the composed series. A property like this forms the basis of the divide and conquer strategy for the fast algorithm as proposed in [4] for the scalar case. Recall also the notation $f^{(V)}(z)$ introduced in the beginning of Section 4 to indicate a truncated series and $\bar{f}^{(V)}(z) = f(z) - f^{(V)}(z)$ the remainder. It will be proved for numerator row degrees, a uniform degree for the denominator and row orders.

Property 6.5 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ be numerator row degrees and row orders and let $U > 0$ be a positive integer. Furthermore, suppose solvability equality (2.4) or (2.5) holds. Let $\tilde{U} = (\tilde{u}_i) \in \mathbf{Z}_+^{p\times 1}$ be defined by*

$$\tilde{u}_i = \begin{cases} v_i - u + 1, & v_i \geq u, \\ 0, & v_i < u. \end{cases}$$

Now we split $f(z)$ as

$$f(z) = f^{(\tilde{U}-1)}(z) + \mathrm{diag}(z^{\tilde{u}_1},\ldots,z^{\tilde{u}_p})g(z)$$

where

$$g(z) = \mathrm{diag}(z^{-\tilde{u}_1},\ldots,z^{-\tilde{u}_p})[f(z)-f^{(\tilde{U}-1)}(z)] = \mathrm{diag}(z^{-\tilde{u}_1},\ldots,z^{-\tilde{u}_p})\bar{f}^{(\tilde{U}-1)}(z).$$

Then the MPAs for f and the MPAs for g are related by

$$[V,U,W;f]_{I,II} = f^{(\tilde{U}-1)}(z) + \mathrm{diag}(z^{\tilde{u}_1},\ldots,z^{\tilde{u}_p})[V-\tilde{U},U,W-\tilde{U};g]_{I,II}.$$

Proof. As in the previous proof we delete the subscript for notational simplification. It is obvious that the solvability equality holds for the problem $(V-\tilde{U},U,W-\tilde{U};g)$. The equality of the two sets will be proved by showing the inclusion in both directions. To show that the right hand side is a subset of the left hand side, take an arbitrary MPA $N_1(z)M_1(z)^{-1} \in [V-\tilde{U},U,W-\tilde{U};g]$. Since it satisfies

$$g(z)M_1(z) - N_1(z) \in \mathbf{E}_{W-\tilde{U}}^{p\times m},$$

we get the order relation

$$\mathrm{diag}(z^{\tilde{u}_1},\ldots,z^{\tilde{u}_p})[g(z)M_1(z)-N_1(z)] \in \mathbf{E}_W^{p\times m}.$$

Therefore, using the definition of $g(z)$ gives

$$f(z)M_1(z) - [f^{(\tilde{U}-1)}(z)M_1(z) + \mathrm{diag}(z^{\tilde{u}_1},\ldots,z^{\tilde{u}_p})N_1(z)] \in \mathbf{E}_W^{p\times m}.$$

By the definition of $\tilde{U}$, we get an upper bound V for the row degrees of the following polynomial matrix

$$f^{(\tilde{U}-1)}(z)M_1(z) + \mathrm{diag}(z^{\tilde{u}_1},\ldots,z^{\tilde{u}_p})N_1(z) \in \mathbf{H}_V^{p\times m}.$$

Therefore, because the orders and the degrees are as they should, we proved the inclusion

$$f^{(\tilde{U}-1)}(z) + \operatorname{diag}(z^{\tilde{u}_1}, \ldots, z^{\tilde{u}_p})[V - \tilde{U}, U, W - \tilde{U}; g] \subset [V, U, W; f].$$

This ends the proof for the inclusion in one direction.

For the inverse inclusion of the above relation, assume now that we have a MPA for f : $N(z)M(z)^{-1} \in [V, U, W; f]$. Like in the first part, we check the order of the approximant first. Since

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m},$$

also

$$[f(z) - f^{(\tilde{U}-1)}(z)]M(z) - [N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{E}_W^{p\times m}.$$

Then scaling with $\operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})$ from the left will bring in the $g(z)$ and the numerator :

$$\operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p}) \left[[f(z) - f^{(\tilde{U}-1)}(z)]M(z) - [N(z) - f^{(\tilde{U}-1)}(z)M(z)]\right] \in \mathbf{E}_{W-\tilde{U}}^{p\times m},$$

i.e.,

$$g(z)M(z) - \operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{E}_{W-\tilde{U}}^{p\times m}.$$

Note that

$$N(z) - f^{(\tilde{U}-1)}(z)M(z) \in \mathbf{E}_{\tilde{U}-1}^{p\times m} \cap \mathbf{H}_V^{p\times m},$$

so that rescaling with $\operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})$ turns this into a polynomial matrix of degree at most $V - \tilde{U}$, i.e.,

$$\operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)] \in \mathbf{H}_{V-\tilde{U}}^{p\times m}.$$

Thus we have proved that

$$\left[\operatorname{diag}(z^{-\tilde{u}_1}, \ldots, z^{-\tilde{u}_p})[N(z) - f^{(\tilde{U}-1)}(z)M(z)]\right] M(z)^{-1} \in [V - \tilde{U}, U, W - \tilde{U}; g].$$

Hence we get the other inclusion.

$$[V, U, W; f] \subset f^{(\tilde{U}-1)}(z) + \operatorname{diag}(z^{\tilde{u}_1}, \ldots, z^{\tilde{u}_p})[V - \tilde{U}, U, W - \tilde{U}; g].$$

So the proof is completed. □

If we are interested in a "block structure" for a MPA table, whatever this may be, we should be able to answer the question : Given some MPA of f for certain V, U, W, is it also a MPA of f for other values of V, U, W? In fact we should ask this question for the complete sets $[V, U, W; f]_{I,II}$. To this we do not have an answer. We can try to solve a simpler problem. From the discussions of Sections 2–3, we remember that the problem $(V, U, W; f)_{I,II}$ can be decoupled column-wise into m sub-problems, i.e., each column of $N(z)$ and $M(z)$ can be determined independently. Now we denote the subproblem for the j-th column by $(V_j, U_j, W_j; f)^j_{I,II}$, and the set of its solutions (N_j, M_j) by $[V_j, U_j, W_j; f]^j_{I,II}$, where V_j, U_j and W_j are the j-th columns of V, U and W respectively. Note however that here the solution set is a collection of column polynomial pairs rather than a collection of rational functions. The following property formulates a simple observation which tells us for type I problems when we have such a pair of numerator-denominator columns for a certain V, U, W, to what other V', U', W' it will be a solution too.

Property 6.6 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$ be such that*

$$\sum_{i=1}^{p} w_i = \sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i.$$

Let $(N(z), M(z))$ be a pair of vectors which solves $(V, U, W; f)_I^j$ i.e., $(N(z), M(z)) \in [V, U, W; f]_I^j$. Let V^* and U^* be the degree vectors of $N(z)$ and $M(z)$ respectively, which are of course bounded by V and U. Furthermore, let W^* be the highest order of the vector $fM - N$. Then for any $V', W' \in \mathbf{Z}_+^{p\times 1}$ and $U' \in \mathbf{Z}_+^{m\times 1}$ satifying the relations

$$V' \geq V^*, \quad U' \geq U^*,$$

$$W' \leq W^*, \quad V' \leq W',$$

and

$$\sum_{i=1}^{p} w'_i = \sum_{i=1}^{p} v'_i + \sum_{i=1}^{m} u'_i,$$

we have

$$(N(z), M(z)) \in [V', U', W'; f]_I^j.$$

The proof of the property is simple and we shall ommit it here. □

This property implies that there is something like a block property of the MPA table, but here, in the matrix case, it is not easy to characterize the structure of the blocks because

(i) The MPA table is multi-dimensional.

(ii) The set $[V, U, W; f]_I^j$ may have many elements (see §8). For the different elements in this set, the degrees V^*, U^* and the orders W^* may be different. So the block may have many "levels".

We should however mention that for the vector case something can be done (see [11]).

We close this section with one final property for these numerator-denominator column pairs. It says that if there exists a linear dependency modulo $z^{v_{\max}}$ between the rows if $f(z)$, where $v_{\max}$ is the maximal degree of a denominator column, then the corresponding numerator column will be zero.

Property 6.7 *Let $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$ satisfy*

$$\sum_{i=1}^{p} w_i = \sum_{i=1}^{p} v_i + \sum_{i=1}^{m} u_i.$$

Let $(N(z), M(z)) \in [V, U, W; f]_I^j$, and

$$\|V\| = \max_{1\leq i\leq p} v_i \leq \min_{1\leq i\leq p} w_i.$$

If for some $A \in \mathbf{C}^{1\times p}$, it happens that

$$Af \in \mathbf{E}_{\|V\|}^{1\times m}, \tag{6.55}$$

then $AN = 0$.

The proof of the property is simple and ommitted □

7 The Duality of MPA

In the case of Padé approximants in a non commutative algebra like e.g., the square matrix case ($p = m$) with uniform degrees and orders (case D1), the equality of left and right approximants is relatively easy to prove. If we think of the vector case ($m = 1$), then only right approximants are relevant since they give a scalar denominator. The corresponding left approximants require a $p \times p$ denominator and are therefore of much less interest. This vector case however illustrates that in the general rectangular matrix case this uniqueness issue is certainly not a trivial matter. This section is devoted to the duality between left and right approximants. Duality means basically that if both solutions of the dual problems exist, then they will be equal. In the next section we shall give results on uniqueness if we assume only the existence of the left or the right approximant.

The duality considered in this section has to be understood in the following sense : To a given right MPA problem, we want to associate a left MPA problem which has the same solutions as the right one. The left problem will be called the dual problem of the right one and conversely. The precise definition is given below.

Definition 7.1 (dual type I problems) *Consider the right MPA problem of type I :* $^R(V, U, W; f)_I$ *and the left problem* $^L(\tilde{V}, \tilde{U}, \tilde{W}; f)_I$. *If for any* $(N, M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m}$ *and* $(\tilde{N}, \tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}}^{p\times p}$ *satisfying*

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}, \tag{7.56}$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) \in \mathbf{E}_{\tilde{W}}^{p\times m}, \tag{7.57}$$

one has

$$\tilde{N}(z)M(z) = \tilde{M}(z)N(z), \tag{7.58}$$

then we say that these first type left and right problems (and their solution sets) are each others dual.

The type II duality is a bit more complicated. It goes as follows :

Definition 7.2 (dual type II problems) *Consider the right and left type II problems* $^R(V, U, W; f)_{II}$ *and* $^L(\tilde{V}, \tilde{U}, \tilde{W}; f)_{II}$. *Define* $U' = (u'_{ij}) \in \mathbf{Z}_+^{m\times m}$ *and* $\tilde{U}' = (\tilde{u}'_{ij}) \in \mathbf{Z}_+^{p\times p}$ *by*

$$u'_{ij} = \min\{u_i, u_j\}, \quad \tilde{u}'_{ij} = \min\{\tilde{u}_i, \tilde{u}_j\}.$$

If for any $(N, M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_{U'}^{m\times m}$ *and* $(\tilde{N}, \tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}'}^{p\times p}$ *satisfying*

$$f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m},$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) \in \mathbf{E}_{\tilde{W}}^{p\times m},$$

one has

$$\tilde{N}(z)M(z) = \tilde{M}(z)N(z),$$

then we say that these second type left and right problems (and their solution sets) are dual.

Similarly, we can also define mixed dualility, i.e., when the left and right problems are of different type. We shall only give a short comment on this in the last Corollary of this section.

First we shall consider the duality of the type I MPAs, i.e., the duality of definition 7.1. The study of this problem shall need the results of the following long and technical Lemma, which we shall prove first. It gives necessary and sufficient conditions for the degrees of the entries of the members of (7.58) to be less than the smallest order involved in (7.56) and (7.57).

Lemma 7.1 *Suppose* $V, W \in \mathbf{Z}_+^{p\times m}$, $U \in \mathbf{Z}_+^{m\times m}$ *and also* $\tilde{V}, \tilde{W} \in \mathbf{Z}_+^{p\times m}$ *and* $\tilde{U} \in \mathbf{Z}_+^{p\times p}$. *Then, under the conditions (compare with the solvability equality (2.4))*

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{k=1}^{m} u_{kj}, \quad j = 1, 2, \ldots, m, \tag{7.59}$$

and

$$\sum_{k=1}^{m} \tilde{w}_{ik} = \sum_{k=1}^{m} \tilde{v}_{ik} + \sum_{k=1}^{p} \tilde{u}_{ik}, \quad i = 1, 2, \ldots, p, \tag{7.60}$$

we have, for $i = 1, \ldots, p$ *and* $j = 1, \ldots, m$,

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_{ik} + u_{kj}\}, \max_{1\le k\le p}\{\tilde{u}_{ik} + v_{kj}\}\right\} \le \min\left\{\min_{1\le k\le m}\{\tilde{w}_{ik}\}, \min_{1\le k\le p}\{w_{kj}\}\right\}, \tag{7.61}$$

if and only if

1. $w_{ij} = \tilde{w}_{st} = w$, *for any* (i, j) *and* (s, t),
2. $u_{ij} = u_i, \quad j = 1, 2, \ldots, m$,
3. $v_{ij} = v_i, \quad j = 1, 2, \ldots, m$,
4. $\tilde{u}_{ij} = \tilde{u}_j, \quad i = 1, 2, \ldots, p$,
5. $\tilde{v}_{ij} = \tilde{v}_j, \quad i = 1, 2, \ldots, p$,
6. $u_i + \tilde{v}_i = w, \quad i = 1, 2, \ldots, m$,
7. $v_i + \tilde{u}_i = w, \quad i = 1, 2, \ldots, p$.

Proof. The proof of sufficiency is simple. We only prove the necessity part.
(a) Suppose for some fixed (i, j) it holds that

$$\min_{1\le k\le m}\{\tilde{w}_{ik}\} \ge \min_{1\le k\le p}\{w_{kj}\}. \tag{7.62}$$

Then (7.61) is equivalent to the following two inequalities

$$\max_{1\le k\le m}\{\tilde{v}_{ik} + u_{kj}\} \le \min_{1\le k\le p}\{w_{kj}\}, \tag{7.63}$$

$$\max_{1\le k\le p}\{\tilde{u}_{ik} + v_{kj}\} \le \min_{1\le k\le p}\{w_{kj}\}. \tag{7.64}$$

We can even decompose this further into the component-wise inequalities

$$\tilde{v}_{ik} + u_{kj} \le \min_{1\le k\le p}\{w_{kj}\}, \quad k = 1, 2, \ldots, m,$$

$$\tilde{u}_{ik} + v_{kj} \le \min_{1\le k\le p}\{w_{kj}\}, \quad k = 1, 2, \ldots, p.$$

Adding these with respect to k gives

$$\sum_{k=1}^{m}(\tilde{v}_{ik} + u_{kj}) \le m \min_{1\le k\le p}\{w_{kj}\}, \tag{7.65}$$

$$\sum_{k=1}^{p}(\tilde{u}_{ik} + v_{kj}) \le p \min_{1\le k\le p}\{w_{kj}\}. \tag{7.66}$$

Using in the sum of (7.59) and (7.60) the previous inequality (7.66) to bound the right hand side, we get

$$\begin{aligned}\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj} &\leq \sum_{k=1}^{m}\tilde{v}_{ik}+\sum_{k=1}^{m}u_{kj}+p\min_{1\leq k\leq p}\{w_{kj}\}\\ &= \sum_{k=1}^{m}(\tilde{v}_{ik}+u_{kj})+p\min_{1\leq k\leq p}\{w_{kj}\}. \qquad (7.11)\end{aligned}$$

With the other inequality (7.65), we can bound this further as

$$\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj}\leq (m+p)\min_{1\leq k\leq p}\{w_{kj}\}. \qquad (7.12)$$

Therefore we have from (7.62) that

$$\tilde{w}_{ik}=\min_{1\leq k\leq p}\{w_{kj}\}=w,\ \ k=1,2,\ldots,m, \qquad (7.13)$$

$$w_{kj}=\min_{1\leq k\leq p}\{w_{kj}\}=w,\ \ k=1,2,\ldots,p. \qquad (7.14)$$

(b) If for fixed (i,j) the opposite of (7.62) holds

$$\min_{1\leq k\leq m}\{\tilde{w}_{ik}\}<\min_{1\leq k\leq p}\{w_{kj}\}, \qquad (7.15)$$

then we can repeat the previous train of deductions with $\min_{1\leq k\leq p}\{w_{kj}\}$ replaced by $\min_{1\leq k\leq m}\{\tilde{w}_{ik}\}$ to find the analog of (7.12)

$$\sum_{k=1}^{m}\tilde{w}_{ik}+\sum_{k=1}^{p}w_{kj}\leq (m+p)\min_{1\leq k\leq m}\{\tilde{w}_{ik}\}.$$

This contradicts (7.15). Thus (7.62) is true for all (i,j). This in turn implies that (7.13) and (7.14) are valid for all (i,j). This proves that relation 1 of the lemma holds.

It follows from (7.63) that

$$\tilde{v}_{ik}+u_{kj}\leq w,\quad \forall(i,j,k),$$

while (7.11) implies

$$mw\leq\sum_{k=1}^{m}(\tilde{v}_{ik}+u_{kj}).$$

Hence

$$\tilde{v}_{ik}+u_{kj}=w,\quad \forall(i,j,k). \qquad (7.16)$$

These relations imply the relations 2, 5 and 6 of the lemma.

Summing up (7.59) and (7.60) and then using (7.16), we get

$$(p+m)w=mw+\sum_{k=1}^{p}(v_{kj}+\tilde{u}_{ik})$$

and thus

$$\sum_{k=1}^{p}(v_{kj}+\tilde{u}_{ik})=pw,\quad \forall(i,j). \qquad (7.17)$$

It follows from (7.61) that

$$v_{kj}+\tilde{u}_{ik}\leq w,\quad \forall(i,j,k). \qquad (7.18)$$

By (7.17) one then may conclude that

$$v_{kj} + \tilde{u}_{ik} = w, \quad \forall (i,j,k).$$

Relations 3, 4 and 7 follow from these equalities. □

We are now ready to state the following duality property for type I MPAs.

Theorem 7.2 *Assume $V, W \in \mathbf{Z}_+^{p\times m}$, $U \in \mathbf{Z}_+^{m\times m}$, $\tilde{V}, \tilde{W} \in \mathbf{Z}_+^{p\times m}$ and $\tilde{U} \in \mathbf{Z}_+^{p\times p}$. If relations 1–7 of Lemma 7.1 hold, then the right MPA problem ${}^R(V,U,W;f)_I$ and the left MPA problem ${}^L(\tilde{V},\tilde{U},\tilde{W};f)_I$ are dual problems in the sense of Definition 7.1.*

Proof. Assume $(N, M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m}$ and $(\tilde{N}, \tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}}^{p\times p}$ satisfy

$$f(z)M(z) - N(z) = R(z) \in \mathbf{E}_W^{p\times m}, \tag{7.19}$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) = \tilde{R}(z) \in \mathbf{E}_{\tilde{W}}^{p\times m}. \tag{7.20}$$

Multiplying (7.19) from the left by $\tilde{M}(z)$ and (7.20) from the right by $M(z)$ and then subtracting, we get

$$\tilde{N}(z)M(z) - \tilde{M}(z)N(z) = \tilde{M}(z)R(z) - \tilde{R}(z)M(z). \tag{7.21}$$

The degree of the (i,j)-th element of the left hand side polynomial of the above equality is less than or equal to

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_{ik} + u_{kj}\}, \max_{1\le k\le p}\{\tilde{u}_{ik} + v_{kj}\}\right\}$$

and the order of the (i,j)-th element on the right hand side is greater than or equal to

$$\min\left\{\min_{1\le k\le m}\{\tilde{w}_{ik}\}, \min_{1\le k\le p}\{w_{kj}\}\right\}.$$

It follows from Lemma 7.1 that if relations 1–7 hold,

$$\tilde{N}(z)M(z) - \tilde{M}(z)N(z) = 0.$$

Thus the theorem is proved. □

Now we consider the duality of the second type MPAs.

Assume we have type II problems ${}^R(V,U,W;f)_{II}$ and ${}^L(\tilde{V},\tilde{U},\tilde{W};f)_{II}$. Recall that for type II problems $U = (u_i) \in \mathbf{Z}_+^{m\times 1}$ and $\tilde{U} = (\tilde{u}_j) \in \mathbf{Z}_+^{1\times p}$. For any $(N, M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_{U'}^{m\times m}$ and $(\tilde{N}, \tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}'}^{p\times p}$ (recall the definition of U' and $\tilde{U}'$ from Definition 7.2) satisfying

$$f(z)M(z) - N(z) = R(z) \in \mathbf{E}_W^{p\times m}$$

and

$$\tilde{M}(z)f(z) - \tilde{N}(z) = \tilde{R}(z) \in \mathbf{E}_{\tilde{W}}^{p\times m},$$

we want the equality

$$\tilde{N}(z)M(z) = \tilde{M}(z)N(z),$$

to be true. Similar to (7.21), we have

$$A\tilde{N}(z)M(z)B - A\tilde{M}(z)N(z)B = A\tilde{M}(z)R(z)B - A\tilde{R}(z)M(z)B,$$

where $A \in \mathbf{C}^{p \times p}$ and $B \in \mathbf{C}^{m \times m}$. In order to make the right hand side of above equality zero, it is sufficient to have

$$\max\left\{\max_{1\le k\le m}\{\tilde{v}_k + u_k\}, \max_{1\le k\le p}\{\tilde{u}_k + v_k\}\right\} \le \min_{i,j,p,q}\{\tilde{w}_{ij}, w_{pq}\} \equiv w,$$

where

$$v_k = \max_i v_{ki} \quad \text{and} \quad \tilde{v}_k = \max_i \tilde{v}_{ik}.$$

Since the maximum of all elements on the left hand side is less than w, each of these elements is less than w, i.e.,

$$\max_{1\le k\le p} \tilde{v}_{kj} + u_j \le w, \quad j = 1, 2, \ldots, m,$$

$$\tilde{u}_i + \max_{1\le k\le m} v_{ik} \le w, \quad i = 1, 2, \ldots, p.$$

Summing up the following two solvablity equalities for type II problems (compare with (2.5))

$$\sum_{k=1}^{m} \tilde{w}_{ik} = \sum_{k=1}^{m} \tilde{v}_{ik} + \sum_{\substack{\tilde{u}_k > \tilde{u}_i \\ 1\le k\le p}} \tilde{u}_i + \sum_{\substack{\tilde{u}_i \ge \tilde{u}_k \\ 1\le k\le p}} \tilde{u}_k + \sum_{\substack{\tilde{u}_k > \tilde{u}_i \\ 1\le k< i}} 1, \quad i = 1, 2, \ldots, p,$$

and

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1\le i\le m}} u_i + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1, \quad j = 1, 2, \ldots, m,$$

we get

$$\begin{aligned}
\sum_{k=1}^{m} \tilde{w}_{ik} + \sum_{k=1}^{p} w_{kj} &= \sum_{k=1}^{m} \tilde{v}_{ik} + \sum_{\substack{\tilde{u}_k > \tilde{u}_i \\ 1\le k\le p}} \tilde{u}_i + \sum_{\substack{\tilde{u}_i \ge \tilde{u}_k \\ 1\le k\le p}} \tilde{u}_k + \sum_{\substack{\tilde{u}_k > \tilde{u}_i \\ 1\le k< i}} 1 \\
&\quad + \sum_{k=1}^{p} v_{kj} + \sum_{\substack{u_i > u_j \\ 1\le i\le m}} u_j + \sum_{\substack{u_i \le u_j \\ 1\le i\le m}} u_i + \sum_{\substack{u_i > u_j \\ 1\le i< j}} 1 \\
&\le \sum_{k=1}^{m} \tilde{v}_{ik} + \sum_{k=1}^{p} \tilde{u}_k + \sum_{k=1}^{p} v_{kj} + \sum_{k=1}^{m} u_k \\
&\le \sum_{k=1}^{m} (\max_i \tilde{v}_{ik} + u_k) + \sum_{k=1}^{p} (\max_j v_{kj} + \tilde{u}_k) \\
&\le (p+m)w.
\end{aligned}$$

On the other hand,

$$\sum_{k=1}^{m} \tilde{w}_{ik} + \sum_{k=1}^{p} w_{kj} \ge (p+m)w.$$

Therefore all the inequalities in the estimation above should be equalities. This leads to

a. $w_{ij} = \tilde{w}_{st} = w$, for any (i, j) and (s, t),

b. $u_{ij} = u_i, \quad j = 1, 2, \ldots, m$, and condition (2.7) in Lemma 2.1 is satisfied.

c. $v_{ij} = v_i, \quad j = 1, 2, \ldots, m$,

d. $\tilde{u}_{ij} = \tilde{u}_j, \quad i = 1, 2, \ldots, p$, and condition (2.7) in Lemma 2.1 is satisfied when U is substituted by $\tilde{U}$.

e. $\tilde{v}_{ij} = \tilde{v}_j, \quad i = 1, 2, \ldots, p$,

f. $u_i + \tilde{v}_i = w, \quad i = 1, 2, \ldots, m$,

g. $v_i + \tilde{u}_i = w, \quad i = 1, 2, \ldots, p$.

Therefore we have proved the following result.

Theorem 7.3 *Let* $V, W \in \mathbf{Z}_+^{p\times m}$, $U \in \mathbf{Z}_+^{1\times m}$, *and* $\tilde{V}, \tilde{W} \in \mathbf{Z}_+^{p\times m}$, $\tilde{U} \in \mathbf{Z}_+^{p\times 1}$. *If relations a–g hold then the type II problems* ${}^R(V, U, W; f)_{II}$ *and* ${}^L(\tilde{V}, \tilde{U}, \tilde{W}; f)_{II}$ *are dual to each other.*

From the deductions of Lemma 7.1 and Theorem 7.3, we can get a mixed duality as in the following corollary.

Corollary 7.4 *Let* $V, W \in \mathbf{Z}_+^{p\times m}$, $U \in \mathbf{Z}_+^{m\times m}$, *and* $\tilde{V}, \tilde{W} \in \mathbf{Z}_+^{p\times m}$, $\tilde{U} \in \mathbf{Z}_+^{p\times p}$.

(i) If relation b in the conditions a–g is replaced by relation 2 in theorem 7.2. Then the problems ${}^R(V, U, W; f)_I$ *and* ${}^L(\tilde{V}, \tilde{U}, \tilde{W}; f)_{II}$ *are dual.*

(ii) If relation d in the conditions a–g is replaced by relation 4 in theorem 7.2. Then the problems ${}^R(V, U, W; f)_{II}$ *and* ${}^L(\tilde{V}, \tilde{U}, \tilde{W}; f)_I$ *are dual.*

8 Uniqueness of MPA

We know from (7.58) that if a right MPA problem ${}^R(V, U, W; f)_I$ and its dual left problem have solutions, then both solutions are unique. This is the implication (ii) $\Rightarrow$ (i) of theorem 8.4. This section will mainly answer a question like: If a left or a right solution exists, but not necessarily both, what conclusions can we draw about the uniqueness ?

We begin this section with the discussion of the uniqueness of the solution of the type I problem ${}^R(V, U, W; f)_I$. The type II problem will be considered later. The left MPA problem can be discussed in a parallel manner, so we do not treat it here. It turns out that the uniqueness problem can be solved only in the case where ${}^R(V, U, W; f)_I$ has a dual problem that may or may not have a solution. As we already know, if the dual problem has a solution, the MPA is unique. We shall prove in Theorem 8.4 that this is also a necessary condition. Moreover we shall prove that it is also necessary and sufficient that a solution of the dual problem exists in a weaker sense. Namely if the set ${}^L R(\tilde{V}, \tilde{U}, \tilde{W}; f)$ of (8.22) has an element with $\det \tilde{M} \neq 0$. Finally a necessary and sufficient condition will also be given which does not use the dual problem. It will be expressed in terms of the dimension of a residual space.

If we should have a dual problem, then it follows from Lemma 7.1 that we should consider row degrees and a uniform order. Therefore we let $V \in \mathbf{Z}_+^{p\times 1}$, $U = (u_i)_{i=1}^m \in \mathbf{Z}_+^{m\times 1}$ and $W = w \in \mathbf{Z}_+$, and we assume that relation (2.4) holds. We set further

$$\tilde{V} = (W - U)^T, \quad \tilde{U} = (W - V)^T \text{ and } \tilde{W} = W.$$

Instead of the strict MPA problems, we shall work in this section with some weaker problems whose solutions are given by the sets

$$ {}^L R(\tilde{V}, \tilde{U}, \tilde{W}; f) = \{(\tilde{N}, \tilde{M}) \in \mathbf{H}_{\tilde{V}}^{p\times m} \times \mathbf{H}_{\tilde{U}}^{p\times p} \setminus \{0\} : \tilde{M} f - \tilde{N} \in \mathbf{E}_{\tilde{W}}^{p\times m}\}, \tag{8.22}$$

$$ {}^{R}R(V,U,W;f) = \{(N,M) \in \mathbf{H}_V^{p\times m} \times \mathbf{H}_U^{m\times m} \setminus \{0\} : fM - N \in \mathbf{E}_W^{p\times m}.\} \quad (8.23) $$

The first set is related to a dual problem of the second one. Furthermore, suppose that there exists a right type I MPA $f^* \in {}^{R}[V,U,W;f]_I \neq \emptyset$. Then we shall consider the following residual space associated with this f^*

$$ R^* = {}^{R}R^*(V,U,W;f) = \{f^*M - N : M(0) = I,\ (N,M) \in {}^{R}R(V,U,W;f)\}. \quad (8.24) $$

Before we come to the characterizations of uniqueness given in Theorem 8.4, we need some preparatory work to find the dimension of the space (8.24). Recall the definition of the matrices ${}^{R}H_I$ of (3.19). The next lemma shall give an explicit form for the elements of the set R^* defined in (8.24).

Lemma 8.1 *Suppose the right MPA $f^* \in {}^{R}[V,U,W;f]_I$ has the formal expansion $f^*(z) = \sum_{k=0}^{\infty} c_k^* z^k$. Let R^* be the set related to this f^* as in (8.24). Then any $R \in R^*$ can be expressed as*

$$ R(z) = z^{w+1} \sum_{k=0}^{\infty} Y_k z^k, \quad (8.25) $$

where

$$ Y_k = G_k[\beta_1, \ldots, \beta_m]. $$

The matrices G_k are given by

$$ G_k = \begin{bmatrix} T_{1,u_1}^{w+k}(f_{11}^*) & T_{1,u_2}^{w+k}(f_{12}^*) & \cdots & T_{1,u_m}^{w+k}(f_{1m}^*) \\ \cdots & \cdots & \cdots & \cdots \\ T_{1,u_1}^{w+k}(f_{p1}^*) & T_{1,u_2}^{w+k}(f_{p2}^*) & \cdots & T_{1,u_m}^{w+k}(f_{pm}^*) \end{bmatrix} \quad (8.26) $$

and the β_i for $i = 1, 2, \ldots, m$ are vectors of dimension $|u| = \sum_{j=1}^{m} u_j$ that are all solutions of

$$ {}^{R}H_I \beta_i = 0, \quad i = 1, \ldots, m, \quad (8.27) $$

with the matrix ${}^{R}H_I$ defined by (3.19).

Proof. By definition, the residual $R \in R^*$ can be written as

$$ R(z) = f^*(z)M(z) - N(z), $$

for some N and M. Suppose f^* is given by $f^* = N^* M^{*-1}$. Then we define ΔM and ΔN by

$$ \Delta M = M - M^* = \begin{bmatrix} \sum_{i=1}^{u_1} \beta_{11}^{(i)} z^i & \cdots & \sum_{i=1}^{u_1} \beta_{1m}^{(i)} z^i \\ \cdots & \cdots & \cdots \\ \sum_{i=1}^{u_m} \beta_{m1}^{(i)} z^i & \cdots & \sum_{i=1}^{u_m} \beta_{mm}^{(i)} z^i \end{bmatrix}, $$

and

$$ \Delta N = N - N^*. $$

Because $f^*M^* - N^* = 0$, it follows that

$$ R(z) = f^*(z)\Delta M(z) - \Delta N(z). $$

If we define the β-vectors with the coefficients of the ΔM matrix as follows :

$$ \beta_i = [\beta_{1i}^{(1)}, \ldots, \beta_{1i}^{(u_1)}; \beta_{2i}^{(1)}, \ldots, \beta_{2i}^{(u_2)}; \cdots; \beta_{mi}^{(1)}, \ldots, \beta_{mi}^{(u_m)}]^T $$

then we have indeed that they satisfy

$$ {}^{R}H_I \beta_i = 0, \quad i = 1, \ldots, m $$

which is (8.27). Since $c_i^* = c_i, \quad$ for $i = 0, 1, \ldots, w$, it follows that

$$\begin{aligned} R(z) &= (\sum_{i=0}^{\infty} c_i^* z^i)\Delta M - \Delta N \\ &= z^{w+1} \sum_{k=0}^{\infty} G_k[\beta_1, \ldots, \beta_m] z^k \\ &= z^{w+1} \sum_{k=0}^{\infty} Y_k z^k. \end{aligned}$$

This concludes the proof of the lemma. □

At this moment we should note that

$$G_0 = \begin{bmatrix} T_{1,u_1}^w(f_{11}) & T_{1,u_2}^w(f_{12}) & \cdots & T_{1,u_m}^w(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{1,u_1}^w(f_{p1}) & T_{1,u_2}^w(f_{p2}) & \cdots & T_{1,u_m}^w(f_{pm}) \end{bmatrix}. \tag{8.28}$$

The i-th row of G_0 is just ${}^L B_I^i$ as defined in (3.21), i.e., the right hand side of the defining equations of the i-th row of the left MPA. This fact will be used in the proof of Theorem 8.4.

Lemma 8.2 *Let us denote the null space of a matrix $A \in \mathbf{C}^{p \times s}$ by $\mathbf{N}(A)$, i.e.,*

$$\mathbf{N}(A) = \{x \in \mathbf{C}^s : \ Ax = 0\}.$$

Suppose we have a matrix $A \in \mathbf{C}^{p \times s}$ and a matrix $B \in \mathbf{C}^{q \times s}$. We shall then denote the range of B when restricted to the nullspace of A by

$$\text{range}(B, A) = \{y \in \mathbf{C}^q : \ y = Bx, \ x \in \mathbf{N}(A)\}.$$

Then the dimension of this restricted null space is given by

$$\dim \text{range}(B, A) = \text{rank} \begin{bmatrix} A \\ B \end{bmatrix} - \text{rank}\, A.$$

The proof of this lemma can be found in [20]. □

Now we are ready to prove the following theorem :

Theorem 8.3 *Let the row degrees V, U and the uniform order W be as in the beginning of this section. Assume that the right type I MPA f^* has the expansion $f^*(z) = \sum_{k=0}^{\infty} c_k^* z^k$. Then the set R^* as defined in (8.24) is a linear space whose dimension is given by*

$$\dim \ R^* = m \sum_{k=0}^{\infty} (\text{rank}\, H_{k+1} - \text{rank}\, H_k),$$

where for $E = W - V \in \mathbf{Z}_+^{p \times 1}$, the matrices H_k are defined by

$$H_k = \begin{bmatrix} T_{e_1+k,u_1}^{v_1}(f_{11}^*) & T_{e_1+k,u_2}^{v_1}(f_{12}^*) & \cdots & T_{e_1+k,u_m}^{v_1}(f_{1m}^*) \\ \cdots & \cdots & \cdots & \cdots \\ T_{e_p+k,u_1}^{v_p}(f_{p1}^*) & T_{e_p+k,u_2}^{v_p}(f_{p2}^*) & \cdots & T_{e_p+k,u_m}^{v_p}(f_{pm}^*) \end{bmatrix}. \tag{8.29}$$

Proof. (i) When we compare the definition of H_0 and the definition of the matrix ${}^R H_I$ in (3.19), we realize that they are the same :

$$H_0 = {}^R H_I.$$

With the result of Lemma 8.1 we can describe the set R^* as

$$R^* = \{R(z) = z^{w+1} \sum_{k=0}^{\infty} G_k B z^k : \ B \in \mathbf{C}^{|u| \times m}, H_0 B = 0\}, \tag{8.30}$$

where $|u| = \sum_{i=1}^{m} u_i$. Therefore R^* is a linear space.

(ii). Consider the space

$$S_k = \{Y = G_k B : \ B \in \mathbf{C}^{|u| \times m}, H_k B = 0\},$$

and suppose

$$Y_{k,1}, Y_{k,2}, \ldots, Y_{k,i_k}$$

is the basis for this space S_k (i_k may be zero) and associate with this basis the matrices

$$B_{k,1}, B_{k,2}, \ldots, B_{k,i_k} \in \{B \in \mathbf{C}^{|u| \times m} : \ H_k B = 0\},$$

such that

$$Y_{k,j} = G_k B_{k,j}, \quad j = 1, 2, \ldots, i_k.$$

Then we shall currently prove that for any $R \in R^*$, which can be written as

$$R(z) = z^{w+1} \sum_{k=0}^{\infty} Y_k z^k = z^{w+1} \sum_{k=0}^{\infty} G_k B z^k \ \in R^*$$

there exist constants α_{ij} such that

$$Y_k = G_k \sum_{l=0}^{\infty} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}. \tag{8.31}$$

In fact, we can prove by induction that there exist α_{ij} such that

$$Y_k = G_k \sum_{l=0}^{k} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}. \tag{8.32}$$

For $k = 0$, since Y_{0j} is a base of S_0, there exist α_{0j} such that

$$Y_0 = \sum_{j=1}^{i_0} \alpha_{0j} Y_{0j} = G_0 \sum_{j=1}^{i_0} \alpha_{0j} B_{0j}.$$

Suppose α_{ij} have been determined for $i = 0, 1, \ldots, k-1$. Note here that the rows of H_k are the rows of H_{k-1} extended with the rows of G_{k-1} for any $k = 1, 2, \ldots$. This implies

$$\begin{aligned} Y_k &= Y_k - G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} + G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} \\ &= G_k \Big(B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}\Big) + G_k \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj} \end{aligned}$$

and

$$G_t(B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = Y_t - Y_t = 0, \quad t = 0, 1, \ldots, k-1,$$

so that $H_k(B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = 0$. i.e.,

$$G_k(B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) \in S_k.$$

Thus there exist α_{kj} such that

$$G_k(B - \sum_{l=0}^{k-1} \sum_{j=1}^{i_l} \alpha_{lj} B_{lj}) = \sum_{j=1}^{i_k} \alpha_{kj} Y_{kj} = G_k \sum_{j=1}^{i_k} \alpha_{kj} B_{kj}.$$

Hence (8.32) is proved. Since G_k is formed by some rows of $H_{k'}$ if $k' > k$, we find that

$$G_k B = 0, \quad \text{if} \quad H_{k'} B = 0, \quad \text{and } k' > k.$$

This fact and (8.32) imply (8.31).

On the other hand, it is easy to show that

$$R_{lj}(z) := z^{w+1} \sum_{k=0}^{\infty} G_k B_{lj} z^k \in R^*, \qquad l = 0, 1, \ldots, \infty; j = 1, 2, \ldots, i_l \tag{8.33}$$

are linearly independent elements from R^*. It then follows from (8.31) that $\{R_{lj}(z)\}$ is a basis for the space R^*. Therefore

$$\dim R^* = \sum_{k=0}^{\infty} i_k = \sum_{k=0}^{\infty} \dim S_k.$$

So we need only consider the dimension of the space S_k.

Now we note that

$$X^m = \{[x_1, \ldots, x_m] \in \mathbf{C}^{d\times m} : \ x_i \in X \subset \mathbf{C}^d \quad i = 1, \ldots, m\}$$

is a $(m \cdot \dim X)$-dimensional linear subspace of $\mathbf{C}^{d\times m}$ provided X is a linear subspace of $\mathbf{C}^d$. Therefore, it follows from

$$\begin{aligned} S_k &= \{y : y = G_k x, \ x \in \mathbf{N}(H_k)\}^m \\ &= \{\text{range}(G_k, H_k)\}^m \end{aligned}$$

and Lemma 8.2, that

$$\dim S_k = m[\text{rank } H_{k+1} - \text{rank } H_k].$$

This concludes the proof of theorem 8.3. □

Given the result about the dimension of R^*, can we now draw a conclusion about the number of elements in ${}^R[V, U, W; f]_I$? Let

$$R_l^* = \text{span}\{R_{l1}, \ldots, R_{li_l}\},$$

where R_{lj} are as defined in (8.33). Then

$$R_i^* \cap R_j^* = \{0\}, \quad \text{for} \quad i \neq j,$$

and

$$\dim R_l^* = \dim S_l = m(\operatorname{rank} H_{k+1} - \operatorname{rank} H_k).$$

Since $G_k B_{lj} = 0$ for $k < l$, we can rewrite (8.33) as

$$R_{lj}(z) := z^{w+1} \sum_{k=l}^{\infty} G_k B_{lj} z^k.$$

Let

$$R_i(z) = f^* M_i - N_i = z^{w+1} \sum_{k=l}^{\infty} Y_k^{(i)} z^k, \; i = 1, 2$$

be two different elements from R_l^*. Then by the linear independency of $\{G_l B_{l1}, \ldots, G_l B_{li_l}\}$, we have $Y_l^{(1)} \neq Y_l^{(2)}$. Hence it follows from

$$\begin{aligned} N_1 M_1^{-1} - N_2 M_2^{-1} &= (f^* - N_2 M_2^{-1}) - (f^* - N_1 M_1^{-1}) \\ &= R_2 M_2^{-1} - R_1 M_1^{-1} \\ &= z^{w+1} \left[(\sum_{k=l}^{\infty} Y_k^{(2)} z^k) M_2^{-1} - (\sum_{k=l}^{\infty} Y_k^{(1)} z^k) M_1^{-1} \right], \end{aligned}$$

that $N_1 M_1^{-1} \neq N_2 M_2^{-1}$. This shows that different residuals in R_l^* will imply that also the corresponding MPAs are different. Furthermore it can be proved similarly that if $i \neq j$, the nonzero residuals in R_i^* and R_j^* are also corresponding to different MPAs. These conclusions give us some idea about how many MPAs there are at least in ${}^R[V, U, W; f]_I$.

Now we are all set up to establish the uniqueness theorem for type I MPAs.

Theorem 8.4 *Consider the row degrees $V \in \mathbf{Z}_+^{p \times 1}$ and $U \in \mathbf{Z}_+^{m \times 1}$ and the uniform order $W \in \mathbf{Z}_+$. Suppose that there exists a right type I MPA, i.e., ${}^R[V, U, W; f]_I \neq \emptyset$. Then the following statements are equivalent.*

(i) The right type I MPA ${}^R[V, U, W; f]_I$ is unique.

(ii) There exists a solution of the dual problem i.e., ${}^L[W - U^T, W - V^T, W; f]_I \neq \emptyset$.

(iii) There exists a $(\tilde{N}, \tilde{M}) \in {}^L R(W - U^T, W - V^T, W; f)$ with $\det \tilde{M} \neq 0$.

(iv) For any $(N, M) \in {}^R R(V, U, W; f)$,

$$ {}^R[V, U, W; f]_I M - N = \{0\}. \tag{8.34}$$

Proof. a) (i) $\Rightarrow$ (ii). If the right approximant ${}^R[V, U, W; f]_I$ is unique, then the set $R^* = \{0\}$ and thus is a space of dimension 0. By Theorem 8.3, we then may conclude that

$$\operatorname{rank} H_1 - \operatorname{rank} H_0 = 0.$$

Since $H_0 = {}^R H_I$ and H_1 is an extension of H_0 obtained by adding the rows of G_0 (see (8.28) and (8.29)), the above relation implies that the existence conditions for the dual problem (3.23) hold for $i = 1, 2, \ldots, p$. From this we conclude that the left MPA exists.

b) (ii) $\Rightarrow$ (iii). Suppose $\tilde{M}^{-1}\tilde{N}$ is a solution of the dual problem ${}^L[W - U^T, W - V^T, W; f]_I$. Then by definition (8.22) $(\tilde{N}, \tilde{M}) \in {}^L R(W - U^T, W - V^T, W; f)$ and because by normalisation of type I MPAs $\tilde{M}(0) = I$, we certainly have $\det \tilde{M} \neq 0$.

c) (iii) $\Rightarrow$ (iv). Suppose $(\tilde{N}, \tilde{M}) \in {}^{L}R(W - U^T, W - V^T, W; f)$ and $\det \tilde{M} \neq 0$. Then for any $(N, M) \in {}^{R}R(V, U, W; f)$, we have by the duality that for any right approximant $N^* M^{*-1} \in {}^{R}[V, U, W; f]_I$,

$$\begin{aligned} \tilde{M}(N^* M^{*-1} M - N) &= \tilde{N} M^* M^{*-1} M - \tilde{M} N \\ &= \tilde{N} M - \tilde{N} M \\ &= 0. \end{aligned}$$

Hence (8.34) holds.

d) (iv) $\Rightarrow$ (i). From (8.34) one has $R^* = \{0\}$. It then immediately follows that the right approximant ${}^{R}[V, U, W; f]_I$ is unique. □

We shall now consider the uniqueness problem for the type II MPAs. First we try to use the ideas of Theorem 5.2 to transform the previous result on type I approximants into type II results. This imposes as in Theorem 5.2 certain restrictions on the degrees and orders. So, we assume that $V, W \in \mathbf{Z}_+^{p \times 1}$ and $U \in \mathbf{Z}_+$. Furthermore suppose that solvability equality (2.5) holds. It follows then from Theorem 5.2 that we have the following relation between type I and type II MPAs

$$ {}^{R}[V, U, W; f]_{II} = f^{(W)}(z) - \operatorname{diag}[z^{w_1}, \ldots, z^{w_p}] \; {}^{R}[U - 1, U, U + W - V - 1; g]_I(z^{-1}), \qquad (8.35)$$

where $g(z) = (g_{ij})_{i,j=1}^{p,m}$ and $g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1})$.

Relation (8.35) implies that there is a one to one map between the type II MPA ${}^{R}[V, U, W; f]_{II}$ for f and the type I MPA ${}^{R}[U - 1, U, U + W - V - 1; g]_I$ for g. Therefore the uniqueness problem of the type II approximant ${}^{R}[V, U, W; f]_{II}$ can be solved by considering the uniqueness problem of the type I approximant ${}^{R}[U - 1, U, U + W - V - 1; g]_I$. In order to use Theorem 8.4 for the uniqueness of the latter, we should have

$$U + W - V - 1 \in \mathbf{Z}_+.$$

i.e., $w_i - v_i = c$ (a constant) for $i = 1, 2, \ldots, p$. This means by (2.5) that p should divide mu. With this extra condition, we can now derive from Theorem 8.4 and Theorem 5.2, the following uniqueness result for type II approximants.

Corollary 8.5 *Let $V, W \in \mathbf{Z}_+^{p \times 1}$ and $W - V, U \in \mathbf{Z}_+$. Let*

$$g(z) = (g_{ij})_{i,j=1}^{p,m}, \qquad \textit{with} \qquad g_{ij}(z) = z^{w_i} f_{ij}^{(w_i)}(z^{-1}).$$

Suppose there exists a type II approximant for f, i.e., ${}^{R}[V, U, W; f]_{II} \neq \emptyset$. Then the following statements are equivalent.

(i) The type II approximant ${}^{R}[V, U, W; f]_{II}$ for f is unique.

(ii) The type I approximant ${}^{R}[U - 1, U, U + W - V - 1; g]_I$ for g is unique.

(iii) The dual type I approximant for g has a solution, i.e., ${}^{L}[W - V - 1, W - V, U + W - V - 1]_I^g \neq \emptyset$.

(iv) There exists a polynomial pair $(\tilde{N}, \tilde{M}) \in {}^{L}R(W - V - 1, W - V, U + W - V - 1; g)$ of the weak dual problem for g such that $\det \tilde{M} \neq 0$.

(v) For any polynomial couple $(N, M) \in {}^{R}R(U - 1, U, U + W - V - 1; g)$, which solves the weaker problem for g, it holds that

$$ {}^{R}[U - 1, U, U + W - V - 1; g]_I M - N = \{0\}.$$

Based on the duality principle of mixed type derived in the previous section (see Corollary 7.4), we have an other uniqueness result.

Corollary 8.6 *Let $V \in \mathbf{Z}_+^{p\times 1}$, $W \in \mathbf{Z}_+$, and $U \in \mathbf{Z}_+^{m\times 1}$ satisfy condition (2.7). If both type I and type II approximants exist, i.e.,*

$$^R[V,U,W;f]_I \neq \emptyset \quad \textit{and} \quad ^R[V,U,W;f]_{II} \neq \emptyset,$$

then the uniqueness of the type I approximant $^R[V,U,W;f]_I$ implies the uniqueness of the type II approximant $^R[V,U,W;f]_{II}$

Proof. If the type I approximant $^R[V,U,W;f]_I$ exists uniquely, then by Theorem 8.4 we know that there exists a solution $^L[W-U^T, W-V^T, W; f]_I$ to the dual problem. Since the latter type I problem $^L(W-U^T, W-V^T, W; f)_I$ is just the mixed dual of the type II problem $^R(V,U,W;f)_{II}$, it follows that the type II approximant $^R[V,U,W;f]_{II}$ is unique. □

The condition in this Corollary is slightly more general than the conditions in Corollary 8.5, but here we only get a sufficient condition for the uniqueness. What conclusions can be drawn from the uniqueness of $^R[V,U,W;f]_{II}$ is a problem to be solved. However, if $^R(V,U,W;f)_{II}$ has a dual problem $^L(\tilde{V},\tilde{U},\tilde{W};f)_{II}$, then we have the following results.

Theorem 8.7 *Let $V \in \mathbf{Z}_+^{p\times 1}$, $W \in \mathbf{Z}_+$, $U \in \mathbf{Z}_+^{m\times 1}$ and let conditions (3.31), (3.32) be satisfied. If a right approximant $^R[V,U,W;f]_{II}$ of type II exists, then it is unique if and only if the dual problem has a solution $^L[(W-U)^T,(W-V)^T,W;f]_{II}$.*

Proof. Since only f is involved, we may drop this from the notation. From Theorem 7.3, we know that the problems $^R(V,U,W)_{II}$ and $^L((W-U)^T,(W-V)^T,W)_{II}$ are dual. Hence by the definition of duality, the uniqueness of $^R[V,U,W]_{II}$ follows from the existence of $^L[(W-U)^T,(W-V)^T,W]_{II}$. It remains to prove the inverse implication. Consider the matrix

$$G = \begin{bmatrix} T_{1,u_1}^{v_1}(f_{11}) & T_{1,u_2}^{v_1}(f_{12}) & \cdots & T_{1,u_m}^{v_1}(f_{1m}) \\ \cdots & \cdots & \cdots & \cdots \\ T_{1,u_1}^{v_p}(f_{p1}) & T_{1,u_2}^{v_p}(f_{p2}) & \cdots & T_{1,u_m}^{v_p}(f_{pm}) \end{bmatrix}$$

and recall the definition of $^RH_{II} = {}^RH_{II}(V,U,W)$ in (3.33). First we show that the restricted range

$$\mathrm{range}(G,\ ^RH_{II}) = \{0\}. \tag{8.36}$$

Let $N^*M^{*-1} \in\ ^R[V,U,W;f]_{II}$ be some type II MPA. Then for any couple (N,M) for which NM^{-1} is in the same set we have

$$R(z) := N^*M^{*-1}M - N = N^*M^{*-1}\Delta M - \Delta N, \tag{8.37}$$

where

$$\Delta M = M - M^*, \quad \Delta N = N - N^*.$$

Since M^* and M are $C1$ canonical, we can write

$$M^*(z) = \mathrm{diag}[z^{u_1},\ldots,z^{u_m}]\tilde{M}^*(z^{-1}), \quad \tilde{M}^*(0) = I$$

and similarly or $M(z)$. Hence

$$\begin{aligned} M^*(z)^{-1} &= \tilde{M}^*(z^{-1})^{-1}\,\mathrm{diag}[z^{-u_1},\ldots,z^{-u_m}] \\ &= \sum_{i=0}^{\infty} d_i z^{-i}\,\mathrm{diag}[z^{-u_1},\ldots,z^{-u_m}], \end{aligned}$$

and

$$R(z) = N^* \sum_{i=0}^{\infty} d_i z^{-i} [\tilde{M}(z^{-1}) - \tilde{M}^*(z^{-1})] - \Delta N.$$

It follows from the uniqueness that $R(z) = 0$, thus that

$$\Delta N = N^* \sum_{i=0}^{\infty} d_i z^{-i} [\tilde{M}(z^{-1}) - \tilde{M}^*(z^{-1})].$$

Comparing the leading coefficient (the highest degree is V) of both sides, we find that

$$\begin{bmatrix} T^{v_j}_{1,u_1}(f_{11}) & \cdots & T^{v_j}_{1,u_m}(f_{1m}) \end{bmatrix} X = 0, \quad \text{for} \quad {}^R H_{II} X = 0, \quad j = 1, \ldots, m,$$

and thus (8.36) holds.

From (8.36) and Lemma 8.2, it follows that (we re-introduce the dependency on V, U and W into the notation)

$$\text{rank } {}^R H_{II}(V,U,W) = \text{rank} \begin{bmatrix} G \\ {}^R H_{II}(V,U,W) \end{bmatrix} = \text{rank } {}^R H_{II}(V-1,U,W).$$

Therefore, by Corollary 3.5, the solution ${}^L[(W-U)^T, (W-V)^T, W]_{II}$ to the dual problem exists and this proves the theorem. □

We should mention here that although the three results concerning the uniqueness of the second type MPA seem to be similar in form, each one treats a different case. So they do not overlap.

9 Conclusion

We have considered some possible definitions of matrix Padé approximants for a power series with rectangular matrix coefficients. Given the problems that occur when we try to generalize the scalar Padé approximants to the matrix case, which are :

(a) The non commutativity of the matrix multiplication,

(b) The many possibilities to fix the degrees and the orders,

(c) The many possibilities of normalization for the approximant,

we had to consider left and right approximants (a). In most cases only one of them should be treated, the other one being similar. The uniqueness problem and the duality of left and right problems as treated in sections 7 and 8 did require a simultaneous treatment of both left and right approximants. Usually we only gave the results for the right approximant. It was illustrated in section 3 how the existence conditions for the left approximant can be derived from the corresponding results for the right one.

For the normalization, we have given two possible choices. One choice, leading to a definition of the type I MPAs, required the denominator $M(z)$ to satisfy $M(0) = I$ which is a reasonable choice if one is approximating in the neighbourhood of $z = 0$ and hence one should require that $M(0)$ is nonsingular to avoid difficulties in $z = 0$. In the type II MPAs, we made a different choice which was inspired by the related problem of minimal partial realization and the derived minimal Padé problem. The idea there is to choose for the denominator a certain canonical representation. This leads to conditions on the degrees of its individual elements which follow from imposed row and column degrees together with specific forms for the matrix of leading row and column coefficients. In a sense we may look at this normalization as being conditions on the leading coefficient of the denominator. The two choices we made in this paper and which gave rise to the type I and type II approximants are only two of the many possibilities and they are not meant to be related or "dual" to each other. However, we do find correspondences between them as given in section 5 where we also indicate the relation with minimal PAs.

Finally, the choice of the degrees and orders which we try to attach to each individual entry of a polynomial matrix or a matrix of series is in our opinion pushing our generalization much farther than it has been done before. This also influenced our choice of the normalization in the type II approximants. Of course, in the most general case of individual degrees and orders, it is most difficult to get nice characterizations, properties, existence, etc. for the approximants. The opposite extreme is to consider a uniform order and degree for all the entries. In this case, one has to be careful with non-commutativity, but the problem is almost scalar in nature. This is the matrix case that has been studied most frequently in the literature.

The solvability equalities (2.4) and (2.5) for the type I and II MPAs expressed that the number of free parameters is equal to the number of approximation conditions. This seems a natural thing to expect in the definition of PAs, but since these equalities are part of the definition of MPAs, they require an extra effort in the completion of certain proofs. These equalities also imposed unexpected restrictions on degrees and orders that are allowed for proving some of the results. The technical result of Lemma 2.1 which reappeared regularly during the development, is a typical exponent of this.

The existence properties derived in section 3 are not spectacular and merely express that a certain set of linear systems of equations defining the denominator should have a solution. Therefore the existence conditions take the form of conditions on the rank of certain matrices. However, even in the scalar case, the block structure of a Padé table which shows where PAs do and where they do not exist, is expressed in terms of the table of Toeplitz or Hankel determinants. For the matrix case, we do exactly the same thing. The importance of the results of section 3 lies in the explicit determinant expressions one can obtain for the MPAs as given in section 4. Moreover, they also give, at least in principle a way to compute the MPAs. Because the systems usually have a special structure, it is possible to design special methods to solve them. These methods are communicated in a separate paper [22]. For the minimal matrix Padé approximants, a Euclidean algorithm exists, which is described in [7]. These MmPAs however are only MPAs of type II in certain cases as described and proved in section 5.

The uniqueness of the MPAs is related to the existence of a dual problem (see sections 7 and 8). Since this can only be proved under relatively restrictive conditions. This is also illustrated by the better studied vector case where we have a generic non-uniqueness, unless extra conditions are imposed.

We do not claim to have developed in this paper the best possible definition of matrix Padé approximants, but we think that even with the restrictions of certain choices we have made, to have posed the problem in a very general setting. This paper just aims at giving a framework which is kept as general as possible for studying matrix Padé approximants. We tried to include most of the existing definitions of vector or matrix Padé approximants (most of them defined with a uniform degree and order) but also left enough room for further development.

References

[1] A.C. Antoulas, On recursiveness and related topics in linear systems, IEEE Trans. Autom. Control, AC-31(1986) 1121–1135.

[2] G.A. Baker Jr., P.R. Graves-Morris, *Padé Approximants*, (2 volumes), Addison-Wesley, Reading, Massachusetts, 1981.

[3] S. Basu, N.K. Bose, Matrix Stieltjes series and network models, SIAM J. Math. Anal., 14(1983) 209–222

[4] R.P. Brent, F.G. Gustavson, D.Y.Y. Yun, Fast solution of Toeplitz systems of equations and computation of Padé approximants, J. Algorithms, 1(1980) 259–295

[5] A. Bultheel, Recursive Relations for Block Hankel and Toeplitz systems, Part I–II, J. Comp. Appl. Math., 10(1984) 301–354.

[6] A. Bultheel, Recursive algorithm for the matrix Padé tabe, Math. Comp., 35(1980) 875–892.

[7] A. Bultheel, M. Van Barel, A matrix Euclidean algorithm and the matrix Padé approximation problem, to appear 1989. See also A matrix Euclidean algorithm and matrix Padé approximations, Report TW 104, January 1988, Department of Computer Science, K.U.Leuven.

[8] A. Bultheel, M. Van Barel, Padé techniques for model reduction in linear system theory: a survey, J. Comp. Appl. Math., 14(1986) 401–438.

[9] A. Bultheel, M. Van Barel, Minimal vector Padé approximation, Presented at the extrapolation and rational approximation conference, Sept. 1989, Luminy, France.

[10] A. Bultheel, G. L. Xu, The structure of Minimal Padé Approximation table, Technical Report TW117, November 1988, Department of Computer Science, K.U.Leuven.

[11] M.G. de Bruin, Generalized C-fractions and a multidimensional Padé table, Ph. D. Thesis, Universiteit Amsterdam, 1974.

[12] A. Draux, Bibliography - Index, Report ANO - 145, Université de Sciences et Techniques de Lille, November 1984.

[13] W.B. Gragg, A. Lindquist, On the partial realization problem, Lin. Alg. Appl., 50(1983) 277–319.

[14] P.R. Graves-Morris, Vector valued rational approximants I, Numer. Math., 42(1983) 331–348, II, IMA J. Numer. Anal., 4(1984) 209–224, III, Constr. Approx., 2(1986) 263–289.

[15] T. Kailath, *Linear systems*, Prentice Hall, Englewood Cliffs, N.J., 1980.

[16] M. Van Barel, A.Bultheel, A Minimal partial algorithm for MIMO systems, I–V, 1988, Technical reports TW79, TW91, TW93, TW94, TW100, Department of Computer Science, K.U.Leuven.

[17] M. Van Barel, Nested minimal partial realizations and related matrix rational approximants, Ph. D. Thesis, Dept. Computer Sc., K.U. Leuven, January 1989.

[18] M. Van Barel, A. Bultheel, A new approach to the rational interpolation problem Presented at the extrapolation and rational approximation conference, Sept. 1989, Luminy, France.

[19] J. van Iseghem, Approximants de Padé vectoriels, Ph. D. Thesis, Lille, 1987.

[20] G.L. Xu, J.K.Li, Generalized Matrix Padé Approximants, Approximation Theory and its Applications, to appear.

[21] G.L.Xu, Existence and Uniqueness of Matrix Padé Approximants, J. Comp. Math., to appear 1989.

[22] G.L.Xu, A. Bultheel, Matrix Padé Approximation : recursive computations, J. Comp. Math., submitted 1989.