SIAM J. COMPUT.
Vol. 18, No. 4, pp. 639-657, August 1989

MATRIX PADÉ FRACTIONS AND THEIR COMPUTATION*

GEORGE LABAHN† AND STAN CABAY‡

Abstract. For matrix power series with coefficients over a field, the notion of a matrix power series remainder sequence and its corresponding cofactor sequence are introduced and developed. An algorithm for constructing these sequences is presented.

It is shown that the cofactor sequence yields directly a sequence of Padé fractions for a matrix power series represented as a quotient $B(z)^{-1}A(z)$. When $B(z)^{-1}A(z)$ is normal, the complexity of the algorithm for computing a Padé fraction of type (m, n) is $O(p^3(m+n)^2)$, where p is the order of the matrices $A(z)$ and $B(z)$.

For a power series that are abnormal for a given (m, n), Padé fractions may not exist. However, it is shown that a generalized notion of Padé fraction, the Padé form, which is introduced in this paper, does always exist and can be computed by the algorithm. In the abnormal case, the algorithm can reach a complexity of $O(p^3(m+n)^3)$, depending on the nature of the abnormalities. In the special case of a scalar power series, however, the algorithm complexity is $O((m+n)^2)$, even in the abnormal case.

Key words. matrix Padé fraction, matrix power series, matrix Padé form

AMS(MOS) subject classifications. 41A21, 41A63, 68Q40

1. Introduction.

Let

$$A(z) = \sum_{i=0}^{\infty} a_i z^i, \tag{1.1}$$

where a_i, $i = 0, \cdots$, is a $p \times p$ matrix with coefficients from a field K, be a formal power series. Loosely speaking, a matrix Padé approximant of $A(z)$ is an expression of the form $U(z) \cdot V(z)^{-1}$, or $V(z)^{-1} \cdot U(z)$, where $U(z)$ and $V(z)$ are matrix polynomials of degree at most m and n, respectively, whose expansion agrees with $A(z)$ up to and including the term z^{m+n}.

The definition of a Padé approximant can be made more formal in a variety of ways. For example, Rissanen [17] restricts $V(z)$ to be a scalar polynomial and allows $U(z)$ to be a $p \times q$ matrix. Typically, however, $U(z)$ and $V(z)$ are $p \times p$ polynomial matrices, and $V(z)$ is further restricted by the condition that the constant term, $V(0)$, is invertible (cf., Bose and Basu [2], Bultheel [5], and Starkand [19]). In this paper, we call such approximants matrix Padé fractions, which is consistent with the scalar $(p = 1)$ case (cf., Gragg [12]).

For a particular m and n, however, matrix Padé fractions need not exist. Therefore, in this paper, we introduce the notion of a matrix Padé form, in which the condition of invertibility of $V(0)$ is relaxed. The definition is a generalization of a similar one given for the scalar case (cf., Gragg [12]). It is shown that matrix Padé forms always exist, but that they may not be unique. In general, matrix Padé forms need not have an invertible denominator, $V(z)$. However, for m and n given, by obtaining a basis for all the Padé forms, we are also able to construct a matrix Padé form with an invertible denominator, $V(z)$, in the case that one does exist.

* Received by the editors April 27, 1987; accepted for publication (in revised form) October 18, 1988.

† Department of Computing Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

‡ Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada T6G 2H1. The work of this author was partially supported by the Natural Science and Engineering Research Council Grant A8035.

Padé approximants have many applications in mathematics and in engineering-related disciplines. Applications include numerical computations for special power series such as the Gamma function (cf., Nemeth and Zimanyi [15]); algorithms in the field of numerical analysis (cf., Gragg [12]); triangulation of block Hankel and Toeplitz matrices (cf., Rissanen [18]); solving linear systems of equations with Hankel of Toeplitz coefficient matrices (cf., Rissanen [16]); in digital filtering theory (cf., Bultheel [7] and Brophy and Salazar [4]; and also in linear control theory (cf., Elgerd [11]).

In the one-dimensional case, examples of algorithms that calculate Padé approximants for normal power series (Gragg [12]) include the ε-algorithm of Wynn [21]; the Levinson–Durbin algorithm [10], [14]; and the algorithm of Trench [20]. Examples of algorithms that are successful in the degenerate nonnormal case include those given by Brent, Gustavson, and Yun [3]; Bultheel [6]; Cabay and Choi [8]; and Rissanen [16].

The matrix case parallels the scalar situation in that most algorithms are restricted to normal power series. Algorithms that require the normality condition include those of Bultheel [5], Bose and Basu [2], Starkand [19], and Rissanen [18]. An algorithm that calculates Padé approximants in a nonnormal case is given by Labahn [13]. However, in his algorithm there are still strict conditions that need to be satisfied by the power series before Padé approximants can be calculated.

The primary contribution of this paper is an algorithm, MPADE, for computing matrix Padé forms for a matrix power series. Central to the development of MPADE are the notions of a matrix power series remainder sequence and the corresponding cofactor sequence, which are introduced in § 4. These are generalizations of notions developed by Cabay and Kossowski [9] for power series over an integral domain. The cofactor sequence computed by MPADE yields a sequence of matrix Padé fractions along a specific off-diagonal path of the Padé table for $A(z)$.

Unlike other algorithms, there are no restrictions placed on the power series in order that MPADE succeed. For normal power series, the complexity of MPADE is $O(p^3 \cdot (m+n)^2)$ operations in K. This is the same complexity as some of the algorithms proposed by Bultheel [5], Bose and Basu [2], Starkand [19], and Rissanen [18]. In the abnormal case, the complexity of the algorithm can reach $O(p^3 \cdot (m+n)^3)$ operations in K, depending on the nature of the abnormalities.

2. Matrix Padé forms. Let $A(z)$ and $B(z)$ be formal power series

$$A(z) = \sum_{i=0}^{\infty} a_i z^i, \qquad B(z) = \sum_{i=0}^{\infty} b_i z^i \tag{2.1}$$

with coefficients from the ring of $p \times p$ matrices over some field K. Throughout this paper it is assumed that the leading coefficient, b_0, of $B(z)$ is an invertible matrix. For nonnegative integers m and n, let

$$U(z) = \sum_{i=0}^{m} u_i z^i, \qquad V(z) = \sum_{i=0}^{n} v_i z^i \tag{2.2}$$

denote $p \times p$ matrix polynomials.

DEFINITION 2.1. The pair of matrix polynomials $(U(z), V(z))$ is defined to be a **right matrix Padé form** (RMPFo) of type (m, n) for the pair $(A(z), B(z))$ if

I. $\partial(U(z)) \leqq m, \partial(V(z)) \leqq n$,†

(2.3) II. $A(z) \cdot V(z) + B(z) \cdot U(z) = z^{m+n+1} W(z)$, where $W(z)$ is a formal matrix power series, and

III. The columns of $V(z)$ are linearly independent over the field K. □

The matrix polynomials $U(z)$, $V(z)$, and $W(z)$ are usually called the right numerator, denominator, and residual (all of type (m, n)), respectively.

There is an equivalent definition for a **left matrix Padé form** (LMPFo). Condition II is replaced with an equivalent version with matrix multiplication by $U(z)$ and $V(z)$ being on the left. Condition III is replaced with the condition that the rows, rather than the columns, of the denominator are linearly independent over the base field K.

However, there is a one-to-one correspondence between RMPFo's and LMPFo's. By taking the transposes of the matrices on both sides of (2.3), it follows that

$$V^t(z) \cdot A^t(z) + U^t(z) \cdot B^t(z) = z^{m+n+1} W^t(z). \tag{2.4}$$

The degree and order conditions are identical. It is clear that if $(U(z), V(z))$ is a RMPFo for $(A(z), B(z))$, then $(U^t(z), V^t(z))$ is a LMPFo for $(A^t(z), B^t(z))$. Thus, any algorithm that calculates a right matrix Padé form of a certain type can also be used to calculate the left matrix Padé form of the same type.

For ease of discussion, we use the following notation. For any matrix polynomial

$$U(z) = u_0 + u_1 z + \cdots + u_k z^k, \tag{2.5}$$

we write U (i.e., the same symbol but without the z variable) to mean the $p(k+1)$ by p vector of matrix coefficients

$$U = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_k \end{bmatrix} \tag{2.6}$$

or, equivalently, $U = [u_0^t, u_1^t, \cdots, u_k^t]^t$.

Let

$$S_{m,n} = \begin{bmatrix} a_0 & & & b_0 & & \\ & \ddots & & & \ddots & \\ \vdots & & a_0 & \vdots & & b_0 \\ & & \vdots & & & \vdots \\ a_{m+n} & \cdots & a_m & b_{m+n} & \cdots & b_n \end{bmatrix} \tag{2.7}$$

denote a Sylvester matrix for $A(z)$ and $B(z)$ of type (m, n). Then (2.3) can be written as

$$S_{m,n} \cdot \begin{bmatrix} V \\ U \end{bmatrix} = 0. \tag{2.8}$$

THEOREM 2.2 (Existence of matrix Padé forms). *For any pair of power series* $(A(z), B(z))$ *and any pair of nonzero integers* (m, n), *there exists a* RMPFo *of type* (m, n).

† $\partial(\ )$ denotes the degree of a matrix polynomial. This is the power of the largest nonzero coefficient of the polynomial.

Proof. Let X denote a vector of length $p(m+n+2)$, and consider the homogeneous system of linear equations

$$S_{m,n} \cdot X = 0. \tag{2.9}$$

Because $S_{m,n}$ has $p(m+n+1)$ rows, it follows that (2.9) has at least p linearly independent solutions. Let $[V', U']'$ denote p such solutions arranged by columns. Then $[V', U']'$ satisfies (2.8); consequently $U(z)$ and $V(z)$, determined according to the convention (2.5) and (2.6), satisfy (2.3). Clearly, the pair $(U(z), V(z))$ also satisfies conditions I in Definition 2.1. Finally, b_0 being nonsingular implies that the linear independence of the columns of $[V', U']'$ is equivalent to the linear independence of the columns of $V(z)$. Thus condition III is satisfied. □

From the proof of Theorem 2.2, it follows that if $S_{m,n}$ has maximal rank, then Padé forms are unique up to multiplication of $U(z)$ and $V(z)$ on the right by a nonsingular matrix. On the other hand, if the rank of $S_{m,n}$ is less than maximal, then more than one independent Padé form exists.

Example 2.3. Let $B(z) = -I$ and

$$A(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} z^5 + \cdots. \tag{2.10}$$

With $m = 2$ and $n = 3$, a basis for the solution space of (2.9) is given by the two vectors

$$X_1 = [0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0, 0, 0]' \tag{2.11}$$

and

$$X_2 = [0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0]'. \tag{2.12}$$

Thus,

$$V = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}' \tag{2.13}$$

and

$$U = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}' \tag{2.14}$$

is a solution of (2.8), and the pair $(U(z), V(z))$, where

$$V(z) = \begin{bmatrix} 0 & 0 \\ 1 - z^2 & z - z^3 \end{bmatrix} \quad \text{and} \quad U(z) = \begin{bmatrix} 0 & 0 \\ 1 & z \end{bmatrix} \tag{2.15}$$

is a Padé form of type (2.3) for $(A(z), B(z))$. □

In Example 2.3, note that the columns of $V(z)$ are linearly independent over the field K, but that they are linearly dependent over the ring of polynomials $K[z]$ (i.e., $V(z)$ is singular). Indeed, for this example, a RMPFo $(U(z), V(z))$ of type $(2, 3)$, for which $V(z)$ is nonsingular, cannot be found. The problem occurs because, although the solution space has dimension 2 when considered as a vector space over the field K, it has only dimension 1 when considered as a module over the ring $K[z]$.

We note that having an invertible denominator is highly desirable, since often the purpose of Padé forms is to approximate the infinite power series

$$-(B(z))^{-1} \cdot A(z) \tag{2.16}$$

by the finite rational form

$$U(z) \cdot (V(z))^{-1}, \tag{2.17}$$

where the approximation is to be exact for the first $m+n+1$ terms. When the denominator is singular, we cannot form this rational expression and this limits the usefulness of Padé approximation. For example, a singular denominator gives no information from the poles since every point is a pole in this case.

3. Matrix Padé fractions. One case when the denominator of a RMPFo is invertible is given by

DEFINITION 3.1. A pair $(U(z), V(z))$ of $p\times p$ matrix polynomials is said to be a **right matrix Padé fraction** (RMPFr) of type (m, n) for the pair $(A(z), B(z))$ if

I. $(U(z), V(z))$ is a RMPFo of type (m, n) for $(A(z), B(z))$, and

II. The constant term, $V(0)$, of the denominator is an invertible matrix. □

Condition II ensures that the denominator, $V(z)$ is an invertible matrix polynomial.

As in the case of Padé forms, there is an equivalent definition for a **left matrix Padé fraction** (LMPFr). Also, there is a correspondence between RMPFr for $(A(z), B(z))$ and LMPFr for $(A(z)', B(z)')$. It is interesting to note that a power series may have a matrix Padé fraction on one side but not on the other. In Example 2.3, the power series $A(z)$ does not have a right matrix Padé fraction of type $(2, 3)$, but it does have a left matrix Padé fraction of type $(2, 3)$. When a power series does have both a right and a left matrix Padé fraction of the same type, then the two resulting rational forms are equal (cf., Baker [1]).

The problem with Padé fractions, as mentioned in the previous section, is that they do not always exist. However, let

$$T_{m,n}=\left[\begin{array}{ccc|ccc} a_0 & & & b_0 & & \\ & \ddots & & & \ddots & \\ \vdots & & a_0 & \vdots & & b_0 \\ & & \vdots & & & \vdots \\ a_{m+n-1} & \cdots & a_m & b_{m+n-1} & \cdots & b_n \end{array}\right] \tag{3.1}$$

and define

$$d_{m,n}=\begin{cases} 1, & m=0, \quad n=0, \\ \det(T_{m,n}), & \text{otherwise}. \end{cases} \tag{3.2}$$

Then, a sufficient condition for the existence of a RMPFr is given by the Theorem 3.2.

THEOREM 3.2. *If $d_{m,n}\neq 0$, then every* RMPFo *of type (m, n) is an* RMPFr *of type (m, n). In addition, a* RMPFr *of type (m, n) is unique up to multiplication on the right by a nonsingular $p\times p$ matrix having coefficients from the field K.*

Proof. Equation (2.8) may be written as follows:

$$\begin{bmatrix} 0 & & & b_0 & & \\ a_0 & & & & & \\ & \ddots & & & \ddots & \\ \vdots & & a_0 & \vdots & & \\ & & & & & b_0 \\ & & \vdots & & & \\ a_{m+n-1} & \cdot & a_m & b_{b+n} & \cdot & b_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \\ \\ v_n \\ u_0 \\ \\ u_m \end{bmatrix} = - \begin{bmatrix} a_0 \\ \\ \\ \vdots \\ \\ \\ a_{m+n} \end{bmatrix} \cdot v_0. \tag{3.3}$$

The matrix on the left of (3.3) is nonsingular, since $d_{m,n}\neq 0$ and b_0 is nonsingular. Thus, all the solutions of (3.3) can be obtained by assigning v_0 arbitrarily and solving (3.3) for the remaining components $v_1, \cdots, v_n, u_0, \cdots, u_m$. If v_0 is chosen to be a singular matrix, then the solution obtained by solving (3.3) violates condition III in the definition of Padé form. Thus, in this case, all RMPFo's are RMPFr's.

To show uniqueness, suppose $(U(z), V(z))$ and $(U'(z), V'(z))$ are two RMPFr's for $(A(z), B(z))$. Then, v_0 and v_0' are both nonsingular matrices with coefficients from the field K. Thus, there exists a nonsingular matrix M with coefficients from K satisfying

$$v_0 = v_0' \cdot M. \tag{3.4}$$

It follows from (3.3) that

$$V(z) = V'(z) \cdot M \quad \text{and} \quad U(z) = U'(z) \cdot M, \tag{3.5}$$

and so uniqueness holds. □

In the next section we also require the following theorem.

THEOREM 3.3. *Let $A(z)$ and $B(z)$ be given by (2.1). If m and n are positive integers such that $d_{m,n} \neq 0$, then* RMPFo's *$(P(z), Q(z))$ of type $(m-1, n-1)$ for $(A(z), B(z))$ are unique up to multiplication of $P(z)$ and $Q(z)$ on the right by a nonsingular matrix from K. In addition, the leading term $R(0)$ of the residual in condition* II *for* RMPFo's,

$$A(z) \cdot Q(z) + B(z) \cdot P(z) = z^{m+n-1} R(z), \tag{3.6}$$

is a nonsingular matrix.

Proof. $S_{(m-1),(n-1)}$ can be obtained from $T_{m,n}$ by deleting the last block row (i.e., the last p rows). Since $T_{m,n}$ is of maximal rank $p(m+n)$, then $S_{(m-1),(n-1)}$ has rank $p(m+n-1)$. Consequently, the dimension of the solution space to

$$S_{(m-1),(n-1)} \cdot X = 0 \tag{3.7}$$

is exactly p. Then, $[Q', P']'$ is obtained by collecting by columns a basis for the solution space of (3.7). Clearly, if $[Q'', P'']'$ and $[Q', P']'$ are two such collections, then there exists a nonsingular matrix M from K such that

$$[Q', P']' = [Q'', P'']' \cdot M. \tag{3.8}$$

Thus, $P(z) = P'(z) \cdot M$ and $Q(z) = Q'(z) \cdot M$, proving uniqueness.

To prove the invertibility of $R(0)$ in (3.6), let $r_0 = R(0)$ and suppose that r_0 is a singular $p \times p$ matrix. Then, there is a nonzero $p \times 1$ vector X that satisfies

$$r_0 \cdot X = 0. \tag{3.9}$$

But, from (3.6) and (3.7), it follows that

$$T_{m,n} \cdot \begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r_0 \end{bmatrix}. \tag{3.10}$$

Thus,

$$T_{m,n} \cdot \begin{bmatrix} Q \\ P \end{bmatrix} \cdot X = 0. \tag{3.11}$$

Since the coefficient matrix for the above system is invertible, we deduce that

$$\begin{bmatrix} Q \\ P \end{bmatrix} \cdot X = 0. \tag{3.12}$$

But this contradicts the fact that the columns of $[Q', P']'$ are made up of linearly independent vectors. This implies that r_0 is invertible. □

The fact that Padé forms of type (m, n) and $(m-1, n-1)$ are uniquely determined after suitable normalizations, when $T_{m,n}$ is nonsingular, allows us to prove such properties as argument invariance (cf., Baker [1]) for the Padé forms computed by the algorithm MPADE given in § 5.

4. Matrix power series remainder sequences. We define a **right matrix Padé table** for $(A(z), B(z))$ to be any infinite two-dimensional collection of RMPFo's of type (m, n) for $(A(z), B(z))$ with $m = 0, 1, \cdots$ and $n = 0, 1, \cdots$. It is assumed that there is precisely one entry (i.e., one RMPFo) assigned to each position in the table. From Theorem 2.2, it follows that a right matrix Padé table exists for any given $(A(z), B(z))$. However, the table is not unique, because RMPFo's are not unique. This is unlike the definition of a Padé table for scalar power series (cf., Gragg [12]), since here a Padé table consists of a collection of Padé fractions, which are unique.

A matrix power series pair $(A(z), B(z))$ is said to be **normal** (cf., Bultheel [5]) if $d_{m,n} \neq 0$ for all m, n. For normal power series, it follows from Theorem 3.2 that every entry in the right matrix Padé table is a RMPFr. Consequently, from condition II in Definition 3.1 of RMPFr's a right-matrix Padé table for normal power series may be made unique by insisting that the constant term $V(0)$ in the denominator of any Padé fraction be the identity matrix.

Following the convention used in the scalar case (cf., Gragg [12]), we also define

$$(U(z), V(z)) = (z^m I, 0) \quad \text{for } m \geqq -1,\ n = -1, \tag{4.1}$$

and

$$(U(z), V(z)) = (0, z^n I) \quad \text{for } m = -1,\ n \geqq 0. \tag{4.2}$$

A right matrix Padé table appended with (4.1) and (4.2) is called an **extended right matrix Padé table**. The use of an extended table is strictly for initialization purposes. The entries given by (4.1) and (4.2) are not right matrix Padé forms (indeed, the $(-1, -1)$ entry is not even a matrix polynomial). However, they do satisfy property II of Definition 2.1. For example, for $m \geqq -1$ and $n = -1$, we have that

$$A(z)V(z) + B(z)U(z) = z^{m+n+1} W(z) \tag{4.3}$$

with

$$W(z) = B(z); \tag{4.4}$$

whereas, for $m = -1$ and $n \geqq 0$, we obtain (4.3) with

$$W(z) = A(z). \tag{4.5}$$

Given the power series (2.1) and any nonnegative integers m and n, we introduce a sequence of points

$$(m_0, n_0), (m_1, n_1), (m_2, n_2), \cdots \tag{4.6}$$

in the extended right matrix Padé table by setting

$$(m_0, n_0) = \begin{cases} (m-n-1, -1) & \text{for } m \geqq n \\ (-1, n-m-1) & \text{for } m < n \end{cases} \tag{4.7}$$

and

$$(m_{i+1}, n_{i+1}) = (m_i + s_i, n_i + s_i), \ i = 0, 1, 2, \cdots, \tag{4.8}$$

where $s_i \geqq 1$. Observe that

$$m_i - n_i = m - n, \qquad i = 0, 1, 2, \cdots, \tag{4.9}$$

and consequently the sequence (4.6) lies along the $m-n$ off-diagonal path of the extended right matrix Padé table. In (4.8), the s_i are selected so that

$$d_{m_{i+1},n_{i+1}} \neq 0 \tag{4.10}$$

and

$$d_{(m_i+j),(n_i+j)} = 0, \tag{4.11}$$

for $j=1,2,\cdots,s_i-1$.

For $i=1,2,\cdots$, let $(U_i(z),\ V_i(z))$ be the unique RMPFr (cf., Theorem 3.2) of type $(m_i,\ n_i)$ for $(A(z),\ B(z))$. Thus $[V_i^t,\ U_i^t]^t$ satisfies

$$S_{m_i,n_i} \cdot \begin{bmatrix} V_i \\ U_i \end{bmatrix} = 0, \tag{4.12}$$

and, according to (2.3), there exists a matrix power series $W_i(z)$ such that

$$A(z) \cdot V_i(z) + B(z) \cdot U_i(z) = z^{m_i+n_i+1} W_i(z). \tag{4.13}$$

Generalizing the notions of Cabay and Kossowski [9], we introduce the following definition.

DEFINITION 4.1. The sequence

$$\{W_i(z)\}, \qquad i=1,2,\cdots, \tag{4.14}$$

is called the **power series remainder sequence** for the pair $(A(z),\ B(z))$. The sequence of pairs

$$\{(U_i(z),\ V_i(z))\}, \qquad i=1,2,\cdots, \tag{4.15}$$

is called the corresponding **cofactor sequence**. The integer pairs $\{(m_i,\ n_i)\}$ are called **nonsingular nodes** along the $m-n$ off-diagonal path of the extended right matrix Padé table for $(A(z),\ B(z))$. □

We note that each term of a power series remainder sequence is unique up to multiplication on the right by a nonsingular matrix. This is also true for each term of the corresponding cofactor sequence.

Initially, when $m \geqq n$, observe that $m_1 = m-n$ and $n_1 = 0$ (i.e., $s_0 = 1$), because in (3.2) the nonsingularity of b_0 implies that $d_{(m-n),0} \neq 0$. Thus, $V_1(z)$ is some arbitrary nonsingular matrix from K and, using (4.12), $U_1(z)$ can be obtained by solving

$$\begin{bmatrix} b_0 & & \\ \vdots & \ddots & \\ b_{m_1} & \cdots & b_0 \end{bmatrix} U_1 = - \begin{bmatrix} a_0 \\ \vdots \\ a_{m_1} \end{bmatrix} V_1. \tag{4.16}$$

That is, $U_1(z)$ can be obtained by multiplying the first m_1+1 terms of the quotient power series $B^{-1}(z) \cdot A(z)$ on the right by $-V_1(z)$.

Initially, when $m<n$, depending on a_0 there are two cases to consider. The simple case, when $\det(a_0) \neq 0$, yields

$$d_{0,(n-m)} = \det \begin{bmatrix} a_0 & & \\ \vdots & \ddots & \\ a_{n-m-1} & \cdots & a_0 \end{bmatrix} \neq 0. \tag{4.17}$$

Thus, $s_0=1$, $m_1=0$, and $n_1=n-m$. Then, the RMPFr $(U_1(z), V_1(z))$ of type (m_1, n_1) is determined by setting $U_1(z)$ to be an arbitrary nonsingular matrix from K and then solving

$$\begin{bmatrix} a_0 & & \\ \vdots & \ddots & \\ a_{n_1} & \cdots & a_0 \end{bmatrix} V_1 = -\begin{bmatrix} b_0 \\ \vdots \\ b_{n_1} \end{bmatrix} U_1. \tag{4.18}$$

That is, when $m<n$ and $\det(a_0)\neq 0$, $V_1(z)$ can be obtained from the first n_1+1 terms of the quotient power series $A^{-1}(z)\cdot B(z)$ multiplied on the right by $-U_1(z)$.

When $m<n$ and $\det(a_0)=0$, we must first determine the smallest positive integer s_0 (i.e., the smallest $m_1=m_0+s_0$ and $n_1=n_0+s_0$) so that $d_{m_1,n_1}\neq 0$. Notice that here $s_0>1$. Once s_0 has been obtained, then $(U_1(z), V_1(z))$ is obtained by solving

$$S_{m_1,n_1}\cdot\begin{bmatrix} V_1 \\ U_1 \end{bmatrix}=0. \tag{4.19}$$

In § 5, we give an algorithm which computes a RMPFo of type (m, n) for $(A(z), B(z))$ by performing a sequence of the above types of initializations (albeit, each for different power series).

When the power series pair $(A(z), B(z))$ is normal, only the initializations corresponding to (4.16) and (4.18) are required. Thus, for normal power series $s_i=1$ for all $i\geqq 1$, and the algorithm reduces to a sequence of truncated power series divisions.

There are also some nonnormal power series that share this property. For each pair of integers m and n, let $r_{m,n}$ be the rank of the matrix $T_{m,n}$. Then normality is equivalent to

$$r_{m,n}=(m+n)\cdot p \tag{4.20}$$

for all m and n. A matrix power series pair $(A(z), B(z))$ is said to be **nearly normal** (cf., Labahn [13]) if, for all integers m and n,

$$r_{m,n}=k_{m,n}\cdot p \tag{4.21}$$

for some integer $k_{m,n}$. Clearly, every normal power series is also a nearly normal power series. In addition, all scalar power series are nearly normal.

For a nearly normal power series pair $(A(z), B(z))$ it is easy to see that when a_0 is singular, then $a_0=0$. This follows from the observation that the rank of a_0 is just $r_{0,1}$, which, if it is not p, must be zero. Also, if $a_0=\cdots=a_{k-1}=0$ and $a_k\neq 0$, then a_k must be a nonsingular matrix for similar reasons. When $k>m$ this implies that there are no nonsingular nodes along the $m-n$ off-diagonal path before and including the node (m, n). Otherwise, when $k\leqq m$, the initialization (4.19) becomes

$$\left[\begin{array}{ccc|ccc} 0 & & & b_0 & & \\ a_{m_1} & & & \vdots & \ddots & \\ \vdots & \ddots & & \vdots & & b_0 \\ a_{m_1+n_1} & \cdots & a_{m_1} & b_{m_1+n_1} & \cdots & b_{n_1} \end{array}\right]\cdot\begin{bmatrix} V_1 \\ U_1 \end{bmatrix}=0, \tag{4.22}$$

where $s_0=k+1$, $m_1=k$, and $n_1=n-m+k$. Consequently the RMPFr $(U_1(z), V_1(z))$ of type (m_1, n_1) is obtained from (4.22) first by setting $U_1(z)=z^{m_1}\cdot U$, where U is any nonsingular matrix from K. Then, $V_1(z)$ is obtained by multiplying the first n_1+1 terms of the quotient power series $(z^{-m_1}\cdot A(z))^{-1}\cdot B(z)$ on the right by $-U$. Thus, also for nearly normal power series (and therefore also for all scalar power series), all initializations reduce to truncated power series divisions.

Corresponding to the power series remainder sequence, we introduce Definition 4.2.

DEFINITION 4.2. The sequence

$$\{(P_i(z), Q_i(z))\}, \qquad i = 1, 2, \cdots, \tag{4.23}$$

where $(P_i(z), Q_i(z))$ is the $(m_i - 1, n_i - 1)$ entry in the extended matrix Padé table for $(A(z), B(z))$, is called a **predecessor sequence** of the power series remainder sequence. □

The pair $(P_i(z), Q_i(z))$ satisfies the equation

$$A(z) \cdot Q_i(z) + B(z) \cdot P_i(z) = z^{m_i + n_i - 1} \cdot R_i(z). \tag{4.24}$$

THEOREM 4.3. *For $i = 1, 2, \cdots$, the predecessors $(P_i(z), Q_i(z))$ are unique up to right multiplication by a nonsingular matrix from K. In addition, the leading term of the residual, $R_i(0)$, is nonsingular.*

Proof. For $m_i > 0$ and $n_i > 0$, the predecessors are right matrix Padé forms and the result is a direct consequence of Theorem 3.3. Thus, it remains only to show that the result holds when either m_1 or n_1 is zero.

When $n_1 = 0$, from (4.7) and (4.8) $m \geqq n$, and the predecessor node is the $(m - n - 1, -1)$ entry of the extended right matrix Padé table. By (4.1), this entry is given uniquely by

$$(P_1(z), Q_1(z)) = (z^{m-n-1} I, 0). \tag{4.25}$$

From (4.4), the residual $R_1(z)$ is $B(z)$ and the theorem therefore holds for $n_1 = 0$, since $\det(b_0) \neq 0$.

When $m_1 = 0$, then from (4.7) and (4.8) $n_1 = n - m > 0$, and the predecessor node is the $(-1, n - m - 1)$ entry of the extended table. By (4.2), this is uniquely given by

$$(P_1(z), Q_1(z)) = (0, z^{n-m-1} I). \tag{4.26}$$

From (4.5) the residual of this node, $R_1(z)$, is $A(z)$. But by (4.17), $(0, n - m)$ is the first nonsingular node if and only if $\det(a_0) \neq 0$. Hence the leading term of the residual is nonsingular. □

The main result of this section is given in Theorem 4.4.

THEOREM 4.4. *For any positive integer k, $(k - 1, k)$ is a nonsingular node in the Padé table for $(W_i(z), R_i(z))$ if and only if $(m_i + k, n_i + k)$ is a nonsingular node in the Padé table for $(A(z), B(z))$.*

Proof. Let M_{11}, M_{21}, M_{12}, and M_{22} be matrices of dimension $p(n_i + k) \times pk$, $p(m_i + k) \times pk$, $p(n_i + k) \times p(k - 1)$, and $p(m_i + k) \times p(k - 1)$, respectively, defined by

$$M_{11} = \begin{bmatrix} v_0 & & \\ & & \\ v_{n_i} & \ddots & \\ & & \\ & \ddots & v_0 \\ & & \\ & & v_{n_i} \end{bmatrix}, \qquad M_{12} = \begin{bmatrix} 0 & & \\ 0 & & \\ q_0 & & \\ & \ddots & \\ q_{n_i - 1} & & q_0 \\ & \ddots & \\ & & q_{n_i - 1} \end{bmatrix}, \tag{4.27}$$

$$M_{21}=\begin{bmatrix} u_0 & & \\ & & \\ u_{m_i} & \ddots & \\ & \ddots & u_0 \\ & & \\ & & u_{m_i} \end{bmatrix}, \qquad M_{22}=\begin{bmatrix} 0 & & \\ p_{-1} & & \\ p_0 & \ddots & \\ & & p_{-1} \\ p_{m_i-1} & & \\ & \ddots & \\ & & p_{m_i-1} \end{bmatrix}. \tag{4.28}$$

In (4.28), $p_{-1}=0$ except when $m_1=0$ and $n_1=0$, in which case, according to (4.25), $p_{-1}=1$. Let M be

$$M=\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}. \tag{4.29}$$

If we set

$$R_i(z)=\sum_{j=0}^{\infty} r_j z^j, \quad \text{with } \det(r_0)\neq 0, \quad \text{and } W_i(z)=\sum_{j=0}^{\infty} w_j z^j, \tag{4.30}$$

then, from (4.13) and (4.24), it follows that

$$T_{(m_i+k),(n_i+k)}\cdot M=\left[\begin{array}{ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ w_0 & & & r_0 & & \\ & \ddots & & & \ddots & \\ \vdots & & w_0 & \vdots & & r_0 \\ & & \vdots & & & \vdots \\ w_{2k-2} & \cdots & w_{k-1} & r_{2k-2} & \cdots & r_k \end{array}\right] = \begin{bmatrix} \mathbf{0} \\ T'_{(k-1),k} \end{bmatrix}, \tag{4.31}$$

where $\mathbf{0}$ represents a zero matrix of size $p(m_i+n_i+1)\times p(2k-1)$ and

$$T'_{(k-1),k}=\left[\begin{array}{ccc|ccc} w_0 & & & r_0 & & \\ & \ddots & & & \ddots & \\ & & w_0 & & & r_0 \\ w_{2k-2} & \cdots & w_{k-1} & r_{2k-2} & \cdots & r_k \end{array}\right]. \tag{4.32}$$

We are now in a position to prove the theorem. Assume $T_{(m_i+k),(n_i+k)}$ is nonsingular. We show that $T'_{(k-1),k}$ is then also nonsingular. Let

$$X=[X_1,\cdots,X_{2k-1}]' \tag{4.33}$$

be a $p(2k-1)\times 1$ vector that satisfies

$$T'_{(k-1),k}\cdot X=0. \tag{4.34}$$

Since $T_{(m_i+k),(n_i+k)}$ is nonsingular, (4.31) implies

$$M\cdot X=0. \tag{4.35}$$

From (4.35), we then obtain that

$$v_0\cdot X_1=0, \tag{4.36}$$

and consequently $X_1=0$, because v_0 is nonsingular. The first block equation from (4.34) then implies

$$r_0 \cdot X_{k+1}=0. \tag{4.37}$$

Thus, $X_{k+1}=0$ because r_0 is nonsingular. In a similar fashion, it follows that $X_2=0$ and $X_{k+2}=0$. Continuing in this way, we obtain that $X=0$, that is, $T'_{(k-1),k}$ is nonsingular.

Conversely, suppose that $T'_{(k-1),k}$ is nonsingular. Let $X=(X_1,\cdots,X_{2k-1})$ be a $1\times p(2k-1)$ vector, $Y=(Y_1,\cdots,Y_{m_i+n_i})$ a $1\times p(m_i+n_i)$ vector, and Z a $1\times p$ vector. Consider

$$(Z, Y, X)\cdot T_{(m_i+k),(n_i+k)}=0. \tag{4.38}$$

Multiplying both sides of (4.38) on the right by M, and using equation (4.31), it follows that

$$X\cdot T'_{(k-1),k}=0. \tag{4.39}$$

Since $T'_{(k-1),k}$ is nonsingular, then $X=0$. Then, in (4.38), using block columns 2 through n_i+1 and block columns n_i+k+2 through to m_i+n_i+k+1 of $T_{(m_i+k),(n_i+k)}$, we obtain

$$Y\cdot T_{m_i,n_i}=0. \tag{4.40}$$

Since T_{m_i,n_i} is nonsingular, $Y=0$. Finally, block column n_i+k+1 of $T_{(m_i+k),(n_i+k)}$ now yields

$$Z\cdot b_0=0. \tag{4.41}$$

Since b_0 is nonsingular, $Z=0$. Hence, $T_{(m_i+k),(n_i+k)}$ is nonsingular.

Theorem 4.4 allows us to calculate nonsingular nodes of a pair of power series by calculating nonsingular nodes of the residual pair of power series. This gives us an iterative method of calculating nonsingular nodes.

THEOREM 4.5. *The cofactor and predecessor sequences for* $(A(z), B(z))$ *satisfy*

$$\begin{bmatrix} U_{i+1}(z) & P_{i+1}(z) \\ v_{i+1}(z) & Q_{i+1}(z)\end{bmatrix}=\begin{bmatrix} U_i(z) & P_i(z) \\ V_i(z) & Q_i(z)\end{bmatrix}\cdot\begin{bmatrix} I & 0 \\ 0 & z^2\cdot I\end{bmatrix}\cdot\begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z)\end{bmatrix}, \tag{4.42}$$

where $(U'(z), V'(z))$ *is the* RMPFr *of type* (s_i-1, s_i) *for* $(W_i(z), R_i(z))$ *and* $(P'(z), Q'(z))$ *is its predecessor.*

Proof. Since $(U_i(z), V_i(z))$ and $(U_{i+1}(z), V_{i+1}(z))$ are successive elements of the cofactor sequence (4.14), then, according to (4.10) and (4.11), (m_i, n_i) and (m_{i+1}, n_{i+1}) are successive nonsingular nodes along the $m-n$ off-diagonal path of the Padé table for $(A(z), B(z))$. By Theorem 4.4, then s_i is the smallest positive integer for which (s_i-1, s_i) is a nonsingular node in the Padé table for $(W_i(z), R_i(z))$. Accordingly, we can determine $(U'(z), V'(z))$ to be the RMPFr of type (s_1-1, s_i) for $(W_i(z), R_i(z))$ and $(P'(z), Q'(z))$ to be its predecessor.

Let $U(z)$, $V(z)$, $P(z)$, and $Q(z)$ be defined by

$$\begin{bmatrix} U(z) & P(z) \\ V(z) & Q(z)\end{bmatrix}=\begin{bmatrix} U_i(z) & P_i(z) \\ V_i(z) & Q_i(z)\end{bmatrix}\cdot\begin{bmatrix} I & 0 \\ 0 & z^2\cdot I\end{bmatrix}\cdot\begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z)\end{bmatrix}. \tag{4.43}$$

We shall first show that $(U(z), V(z))$ given by (4.43) is the RMPFr of type (m_{i+1}, n_{i+1}) for $(A(z), B(z))$. Because RMPFr's are unique, then $(U(z), V(z))$ must be the $(i+1)$st term in the cofactor sequence.

Since $(U'(z), V'(z))$ is a RMPFr of type (s_i-1, s_i) for $(W_i(z), R_i(z))$, it satisfies

$$W_i(z)\cdot V'(z)+R_i(z)\cdot U'(z)=z^{2s_i}W'(z), \tag{4.44}$$

where $V'(0)$ is nonsingular. Then, using (4.13), (4.24), (4.43), and (4.44), we get

$$\begin{aligned} A(z)\cdot V(z)+B(z)\cdot U(z) &\\ &= A(z)\cdot\{V_i(z)\cdot V'(z)+z^2Q_i(z)\cdot U'(z)\}\\ &\quad + B(z)\cdot\{U_i(z)\cdot V'(z)+z^2P_i(z)\cdot U'(z)\}\\ &= \{A(z)\cdot V_i(z)+B(z)\cdot U_i(z)\}\cdot V'(z)\\ &\quad +\{A(z)\cdot Q_i(z)+B(z)\cdot P_i(z)\}\cdot z^2U'(z)\\ &= z^{m_i+n_i+1}\cdot\{W_i(z)\cdot V'(z)+R_i(z)\cdot U'(z)\}\\ &= z^{(m_i+s_i)+(n_i+s_i)+1}\cdot W'(z). \end{aligned} \tag{4.45}$$

Thus, condition II for a RMPFo of type (m_i+s_i, n_i+s_i) for $(A(z), B(z))$ is satisfied.

To verify condition I, expanding (4.43) gives

$$U(z)=U_i(z)\cdot V'(z)+z^2P_i(z)\cdot U'(z), \tag{4.46}$$

so that

$$\begin{aligned} \partial(U(z)) &\leqq \max(m_i+s_i, 2+(m_i-1)+s_i-1)\\ &= m_i+s_i. \end{aligned} \tag{4.47}$$

Similarly,

$$\partial(V(z))\leqq n_i+s_i. \tag{4.48}$$

Finally, to verify condition II for a RMPFr (and, thus, condition III for a RMPFo, as well), observe that

$$V(0)=V_i(0)\cdot V'(0), \tag{4.49}$$

and, consequently, $V(0)$ is invertible since both $V_i(0)$ and $V'(0)$ are invertible. Notice that we have a somewhat stronger result here, namely that if $V_i(z)$ and $V'(z)$ are both normalized with $V_i(0)=I$ and $V'(0)=I$, then so is $V(z)$.

Therefore, $(U(z), V(z))$ is a RMPFr of type (m_i+s_i, n_i+s_i) for $(A(z), B(z))$. Thus, it is the $(i+1)$st term in the cofactor sequence and $W'(z)$ is the $(i+1)$st term in the power series remainder sequence for $(A(z), B(z))$.

Notice that the above arguments also hold in the special case when $m=n$ and $i=1$. In this case $P_1(z)=z^{-1}I$ which is not a matrix polynomial. However, the right side of equation (4.43) immediately multiplies the predecessor by z^2 which subsequently results in a matrix polynomial.

A similar argument shows that $(P(z), Q(z))$ given in (4.43) is the predecessor of the nonsingular node (m_{i+1}, n_{i+1}). Hence the recurrence relation (4.42) holds. □

For purposes of the algorithm given in the next section, observe that if $(U'(z), V'(z))$ is a RMPFo of type $(s-1, s)$ for $(W_i(z), R_i(z))$ and $(P'(z), Q'(z))=(0, I)$ then in (4.43) $(U(z), V(z))$ yields a RMPFo (rather than a RMPFr) of type (m_i+s, n_i+s) for $(A(z), B(z))$ and $(P(z), Q(z))=(U_i(z), V_i(z))$.

5. The algorithm. Given nonnegative integers m and n, the algorithm MPADE below makes use of Theorem 4.5 to compute the cofactor and predecessor sequences

(4.10) and (4.18), respectively. Thus, intermediate results available from MPADE include those RMPFr's $(U_i(z), V_i(z))$ for $(A(z), B(z))$ at all the nonsingular nodes (m_i, n_i), $i = 1, 2, \cdots, k-1$, smaller than (m, n), along the off-diagonal path $m_i - n_i = m - n$. The output gives results associated with the final node (m_k, n_k). If (m, n) is also a nonsignular node, then the output $(U_k(z), V_k(z))$ is a RMPFr of type (m, n) for $(A(z), B(z))$, and $(P_k(z), Q_k(z))$ is a RMPFo of type $(m-1, n-1)$. If (m, n) is a singular node, then the output $(U_k(z), V_k(z))$ is simply a RMPFo of type (m, n) for $(A(z), B(z))$, and now $(P_k(z), Q_k(z))$ is set to be the RMPFr of type (m_{k-1}, n_{k-1}). An exception occurs in the latter case when $k = 0$ and $m < n$. Here, all nodes along the off-diagonal path must have been singular, and for $(P_k(z), Q_k(z))$ the algorithm returns instead the initial value $(0, z^{n-m-1}I)$.

Note that, when (m, n) is not a nonsingular node, a simple modification of MPADE allows the computation of all RMPFo's of type (m, n) for $(A(z), B(z))$. It is only necessary to arrange to compute q columns of $[V'_k, U'_k]$, rather than p, in order to form a basis for the solution space of the equation in step 3.1 of MPADE. From this basis, it is then possible to construct a $p \times p$ matrix $V(z)$, and a corresponding $U(z)$, for which $(U(z), V(z))$ is a RMPFo of type (m, n) for $(A(z), B(z))$ and has the property that $V(z)$ is an invertible matrix, assuming such a RMPFo exists. This enhancement is not included in MPADE primarily to simplify the presentation of the algorithm.

MPADE(A, B, m, n)

If $m \geqq n$

then

M1) $\begin{bmatrix} m_0 \\ n_0 \end{bmatrix} \leftarrow \begin{bmatrix} m-n-1 \\ -1 \end{bmatrix}$

M2) $\left(s_0, \begin{bmatrix} V_1(z) & Q_1(z) \\ U_1(z) & P_1(z) \end{bmatrix}\right) \leftarrow INITIAL\_PADE(B(z), A(z), n, m)$

else

M3) $\begin{bmatrix} m_0 \\ n_0 \end{bmatrix} \leftarrow \begin{bmatrix} -1 \\ n-m-1 \end{bmatrix}$

M4) $\left(s_0, \begin{bmatrix} U_1(z) & P_1(z) \\ V_1(z) & Q_1(z) \end{bmatrix}\right) \leftarrow INITIAL\_PADE(A(z), B(z), m, n)$

M5) $\begin{bmatrix} m_1 \\ n_1 \end{bmatrix} \leftarrow \begin{bmatrix} m_0 + s_0 \\ n_0 + s_0 \end{bmatrix}$

M6) $i \leftarrow 1$

M7) Do while $m_i < m$

M8) Compute $R_i(z)$ satisfying (4.24)

M9) Compute $W_i(z)$ satisfying (4.13)

M10) $\left(s_i, \begin{bmatrix} U'(z) & P'(z) \\ V'(z) & Q'(z) \end{bmatrix}\right) \leftarrow INITIAL\_PADE(W_i(z), R_i(z), m - m_i - 1, n - n_i)$

M11) $\begin{bmatrix} U_{i+1}(z) & P_{i+1}(z) \\ V_{i+1}(z) & Q_{i+1}(z) \end{bmatrix} \leftarrow \begin{bmatrix} U_i(z) & P_i(z) \\ V_i(z) & Q_i(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix}$

M12) $\begin{bmatrix} m_{i+1} \\ n_{i+1} \end{bmatrix} \leftarrow \begin{bmatrix} m_i + s_i \\ n_i + s_i \end{bmatrix}$

M13) $i \leftarrow i+1$

End do

M14) $k \leftarrow i$

M15) Return $\left(\begin{bmatrix} U_k(z) & P_k(z) \\ V_k(z) & Q_k(z) \end{bmatrix}\right)$

End MPADE

INITIAL_PADE(W(z), R(z), m', n')

I1) $s \leftarrow 0$

I2) $d \leftarrow 0$

I3) Do while $s \leqq m'$ and $d = 0$

I4) Compute $d \leftarrow \det(T_{s,n'-m'+s})$

I5) $s \leftarrow s+1$

End do

I6) Solve

$$S_{s-1,n'-m'+s-1} \cdot \begin{bmatrix} V' \\ U' \end{bmatrix} = 0$$

If $s > 1$ and $d \neq 0$

then

I7) Solve

$$S_{s-2,n'-m'+s-2} \cdot \begin{bmatrix} Q' \\ P' \end{bmatrix} = 0$$

else

I8) $\begin{bmatrix} Q' \\ P' \end{bmatrix} \leftarrow \begin{bmatrix} z^{n'-m'-1} I \\ 0 \end{bmatrix}$

I9) Return $\left(s, \begin{bmatrix} U' & P' \\ V' & Q' \end{bmatrix}\right)$

End INITIAL_PADE

THEOREM 5.1. *The* MPADE *algorithm is valid.*

Proof. The argument is by induction on i.

Initially, in step M2 of MPADE, where $m \geqq n$, the parameters input to INITIAL_PADE are $W(z) = B(z)$, $R(z) = A(z)$, $m' = n$ and $n' = m$. Consequently, INITIAL_PADE computes $s = 1$, since in step I5

$$d = \det(T_{0,m-n}) = \det \begin{bmatrix} b_0 & & \\ & \ddots & \\ b_{m-n-1} & & b_0 \end{bmatrix} \neq 0 \tag{5.1}$$

when $m > n$, and $d = 1$ when $m = n$. In step I6, the algorithm solves

$$S_{0,m-n} \begin{bmatrix} V' \\ U' \end{bmatrix} = \begin{bmatrix} b_0 & & & a_0 \\ & \ddots & & \\ b_{m-n} & & b_0 & a_{m-n} \end{bmatrix} \begin{bmatrix} V' \\ U' \end{bmatrix} = 0 \tag{5.2}$$

and step I8 yields

$$\begin{bmatrix} Q' \\ P' \end{bmatrix} = \begin{bmatrix} z^{m-n-1} I \\ 0 \end{bmatrix}. \tag{5.3}$$

Since the substitution in step M2 of MPADE yields $s_0 = s$ and

$$\begin{bmatrix} U_1 & P_1 \\ V_1 & Q_1 \end{bmatrix} = \begin{bmatrix} V' & Q' \\ U' & P' \end{bmatrix}, \tag{5.4}$$

it follows that the initialization for $m \geqq n$ is exactly that given by (4.16) and (4.25).

Alternately, when $m < n$, initialization is accomplished by step M4. In this case, the parameters input to INITIAL_PADE are $W(z) = A(z)$, $R(z) = B(z)$, $m' = m$, and $n' = n$. If $\det(a_0) \neq 0$, INITIAL_PADE again computes $s = 1$, since in step I5

$$d = \det(T_{0,n-m}) = \det \begin{bmatrix} a_0 & & \\ & \ddots & \\ a_{n-m-1} & & a_0 \end{bmatrix} \neq 0. \tag{5.5}$$

Step I6 then solves (4.18) with $U_1 = U'$ and $V_1 = V'$, and, with the substitution $(P_1, Q_1) = (P', Q')$, step I7 yields the required predecessor (4.26). If $\det(a_0) = 0$, then step I5 determines the smallest integer $s \geqq 2$, if one exists, for which $d = \det(T_{s-1,n-m+s-1}) \neq 0$. Consequently, step I6 solves exactly the system (4.19) and step I7 must then yield the correct predecessor.

Assume that, for $i \geqq 1$, MPADE calculates $(U_i(z), V_i(z))$ and $(P_i(z), Q_i(z))$ correctly. We shall show that one pass through the while loop M7 correctly computes $(U_{i+1}(z), V_{i+1}(z))$ and its predecessor.

In step M10 and MPADE, the parameters input to INITIAL_PADE are $W(z) = W_i(z)$, $R(z) = R_i(z)$, $m' = m - m_i - 1$, and $n' = n - n_i$. Noting (4.9), step I4 computes the smallest positive integer s, if one exists, for which $d = \det(T_{s-1,s}) \neq 0$. Clearly, then I6 computes a RMPFr of type $(s-1, s)$ for $(W_i(z), R_i(z))$, and steps I7 and I8 its predecessor. Thus, the matrix polynomials in step M11 correspond exactly to those of (4.42); that is, the algorithm correctly computes $(U_{i+1}(z), V_{i+1}(z))$ and its predecessor.

To complete the proof of algorithm validity, consideration must be given to the case for which there exists no s such that $d \neq 0$ in the while loop I3 of INITIAL_PADE. On exit from the while loop, observe that $s = m' + 1$. Step I6 then computes $(U'(z), V'(z))$ to be a RMPFo of type (m', n') for $(W(z), R(z))$ and sets $(P'(z), Q'(z)) = (0, z^{n-m-1} I)$.

The case where there exists no s such that $d \neq 0$ can occur when INITIAL_PADE is invoked in steps M4 and M10 of MPADE, only. If it occurs at step M4, then $(U_1(z), V_1(z))$ becomes a RMPFo of type (m, n) for $(A(z), B(z))$ as computed by INITIAL_PADE, and $(P_1(z), Q_{1(z))} = (0, z^{n-m-1} I)$. Since step M5 next yields $(m_1, n_1) = (m, n)$, the algorithm immediately terminates. On the other hand, if it occurs at step M10, then $s_i = m - m_i$, $(U'(z), V'(z))$ is a RMPFo of type $(s_i - 1, s_i)$ for $(W_i(z), R_i(z)$ and $(P'(z), Q'(z)) = (0, I)$. Accordingly (cf., last paragraph of § 4), $(U_{i+1}(z), V_{i+1}(z))$ computed in step M11 is a RMPFo of type (m, n) for $(A(z), B(z))$ and $(P_{i+1}(z), Q_{i+1}(z)) = (U_i(z), V_i(z))$. Since step M12 yields $(m_{i+1}, n_{i+1}) = (m, n)$, the algorithm terminates. □

6. Complexity of the MPADE algorithm. Note that, in steps M8 and M9 of MPADE, only the first $m + n - m_i - n_i$ terms in $R_i(z)$ and $W_i(z)$ are required to ensure the subsequent success of step M10. Indeed, only the first $2s_i$ terms, $s_i \leqq m - m_i$, are sufficient, but unfortunately s_i is not known prior to step M10. Nevertheless, an efficient implementation can take advantage of this observation by delaying the computation of $R_i(z)$ and $W_i(z)$. Declaring $(A(z), B(z))$, $(U_i(z), V_i(z))$, $(P_i(z), Q_i(z))$ to be global variables, the coefficients of $R_i(z)$ and $W_i(z)$ can be computed in INITIAL_PADE only when they become necessary. The cost analysis below assumes that the algorithm has been implemented in such a fashion.

In assessing the costs of MPADE, it is assumed that classical algorithms are used for the multiplication of polynomials. Only the more costly steps are considered. For these steps, Table 6.1 below provides cruder upper bounds on the number of multiplications in K required. The table provides separate bounds for the normal and abnormal cases.

In step I5 of INITIAL_PADE, it is assumed that the Gaussian elimination method is used to obtain the LU decomposition of $T_{(s-1),n'-m'+s-1}$. In addition, it is assumed that Gaussian elimination is accompanied with bordering techniques. Thus, as s increases by 1 in step I4, the results of the previous pass through the while loop are used to achieve the current LU decomposition. The bound for step I5 in Table 6.1 for the abnormal case assumes we do not take any advantage of the special nature of $T_{(s-1),n'-m'+s-1}$. In the normal case $s=1$, and $T_{0,n'-m'}$ is already in triangular form, and so no computation is required in step I5.

For step I6, it is assumed that the LU decomposition of $T_{(s-1),n'-m'+s-1}$ from step I5, is used to simplify the triangulation of $S_{(s-1),n'-m'+s-1}$. The solution $[V', U']$ is obtained finally by solving this triangularized S. Similar observations apply to step I7 in the abnormal case.

Since steps M2, M4, and M10 simply invoke INITIAL_PADE, estimates of their costs are obtained by summing the costs of steps I5, I6, and I7 with appropriate substitutions of variables. Note that the cost of M2 is the same in both the normal and abnormal case, since for $m \geqq n$ it is always true that $s_0 = 1$.

An upper bound for the number of multiplications in K required by MPADE is obtained by summing the costs of the last six rows in Table 6.1 for $i = 0, 1, \cdots, k$. We use the fact that

$$\sum_{i=0}^{k} s_i = m, \text{ if } m \geqq n, \text{ and } \sum_{i=0}^{k} s_i = n, \text{ if } m < n. \tag{6.1}$$

In addition,

$$\sum_{i=0}^{k} m_i^{\alpha} s_i^{\beta} \leqq m^{\alpha+\beta} \text{ and } \sum_{i=0}^{k} n_i^{\alpha} s_i^{\beta} \leqq n^{\alpha+\beta}. \tag{6.2}$$

Then, steps M4 and M10 in the abnormal case have a complexity of $O(p^3(m+n)^3)$ and the remaining steps a complexity of $O(p^3(m+n)^2)$, at worst. When $(A(z), B(z))$ is normal, then due to the fact that $T_{s-1,n'-m'+s-1}$ is always in triangular form, the complexity of MPADE reduces to $O(p^3(m+n)^2)$. This is also true when $(A(z), B(z))$ is nearly normal. In this case s_i is often larger than 1, but the matrix $T_{s-1,n'-m'+s-1}$ is

TABLE 6.1
Bounds on operations per step.

Step	Normal case	Abnormal case
I5	0	$p^3(n'-m'-1+2(s-1))^3/3$
I6	$p^3(n'-m'+1)^2/2$	$p^3(n'-m'+2s-1)^2/2$
I7	0	$p^3(n'-m'+2s-1)^2/2$
M2	$p^3(m-n)^2$	$p^3(m-n)^2$
M4	$p^3(m-n)^2$	$p^3[(n-m-1+2(s_0-1))^3/3+(n-m+2s_0-1)^2]$
M8	$2p^3(m_i+n_i+2)$	$2p^3(m_i+n_i+2)s_i$
M9	$2p^3(m_i+n_i+2)$	$2p^3(m_i+n_i+2)s_i$
M10	0	$8p^3(s_i-1)^3/3$
M11	$8p^3(m_i+n_i+2)$	$8p^3(m_i+n_i+2)s_i$

also always in triangular form and so again the complexity is $O(p^3(m+n)^2)$. In particular, in the scalar case the complexity of MPADE is $O((m+n)^2)$.

The algorithm gives the worst performance when no nonsingular nodes are encountered along the $m-n$ off-diagonal path. In this case, with $m<n$, the algorithm reduces to solving one Sylvester system

$$S_{m,n}\begin{bmatrix} V_1 \\ U_1 \end{bmatrix} = 0 \tag{6.3}$$

in step M4 of MPADE. In Table 6.1, with $s_0 = m+1$, then the cost is simply that of Gaussian elimination, namely, approximately $p^3(m+n)^3/3$. Note that with the existence of even one nonsingular node the cost of MPADE can be dramatically reduced. If, for example, this one nonsingular node is $(m_1, n_1) = (m/2, n-m/2)$, where m is even, then $s_0 = 1+m/2$, $s_1 = m/2$ and the algorithm reduces essentially to solving (in steps M4 and M10) two Sylvester systems, each of approximately half the total size. This results in a saving of a factor of 4 over the simple use of Gaussian elimination. Algorithms requiring normality, on the other hand, break down when even one intermediate node is singular.

7. Conclusions. We have considered the problem of determining an adequate definition for a rational approximant of a formal matrix power series and also, given a suitable definition, the problem of computing it. We have restricted our attention to square matrix power series.

In attempting to extend the notion of Padé approximation to matrix power series, we have followed the classical theory of Padé approximants for scalar power series. We introduce the notion of a Padé form, which always exist but may not be unique, and also the notion of Padé fraction, which is unique but need not exist. The definition of Padé form is meant to be as broad as possible. By constructing all the Padé forms of type (m, n), it is always possible to determine ones for which the denominator is invertible, should one exist.

The notion of a matrix power series remainder sequence introduced in this paper is a generalization of one given by Cabay and Kossowski [9] for scalar power series. The cofactor sequence, which is shown to be associated with the remainder sequence, yields directly all the Padé fractions at the nonsingular nodes of a particular off-diagonal path of the Padé table. By determining also the (unique) Padé form at nodes preceding the nonsingular nodes, we are able to compute Padé fractions iteratively from one nonsingular node to the next. The resulting algorithm is at least as fast as other algorithms for computing matrix Padé fractions, and it is the only one that succeeds in the abnormal case.

The algorithm can be improved in a number of ways. We expect that the cost of the decomposition of $T_{s-1,n'-m'+s-1}$ in step I5 and, consequently, of $S_{s-1,n'-m'+s-1}$ in step I6 can be improved by taking advantage of the special structure of Sylvester matrices. The algorithm would also experience an improvement if it were possible to identify additional points between nonsingular nodes for which Theorem 4.5 is valid. This would improve the algorithm by decreasing the s_i. This, and in general the nature of Padé forms between nonsingular nodes, is a subject for further research. Finally, by appealing to fast methods for polynomial arithmetic, it is of interest to attempt to develop a recursive divide-and-conquer version of MPADE.

For normal and nearly normal power series, progressing from one nonsingular node to the next is equivalent to power series division of the residuals associated with the nonsingular nodes (because $S_{s-1,n'-m'+s-1}$ in step I5 of INITIAL_PADE, with the exception of one column, reduces to a triangular matrix). Thus, in this case and in addition when $A(z)$ and $B(z)$ are matrix polynomials, there is a strong analogy between MPADE and Euclid's algorithm. It is a subject for future research to investigate the possibility of using MPADE to compute the greatest common divisor of two matrix polynomials in the abnormal case.

REFERENCES

[1] G. A. BAKER, *Essentials of Padé Approximants*, Academic Press, New York, 1975.

[2] N. K. BOSE AND S. BASU, *Theory and recursive computation of* 1-D *matrix Padé approximants*, IEEE Trans. on Circuits and Systems, 4 (1980), pp. 323-325.

[3] R. BRENT, F. G. GUSTAVSON, AND D. Y. Y. YUN, *Fast solution of Toeplitz systems of equations and computation of Padé approximants*, J. Algorithms, 1 (1980), pp. 259-295.

[4] F. BROPHY AND A. C. SALAZAR, *Considerations of the Padé approximant technique in the synthesis of recursive digital filters*, IEEE Trans. Audio and Electroacoustics, AU-21 (1973), pp. 500-505.

[5] A. BULTHEEL, *Recursive algorithms for the matrix Padé table*, Math. Computation, 35 (1980), pp. 875-892.

[6] ———, *Recursive algorithms for nonnormal Padé tables*, SIAM J. Appl. Math., 39 (1980), pp. 106-118.

[7] ———, *Algorithms to compute the reflection coefficients of digital filters*, Numer. Methods Approx. Theory, 7 (1983), pp. 33-50.

[8] S. CABAY AND D. K. CHOI, *Algebraic computations of scaled Padé fractions*, SIAM J. Comput., 15 (1986), pp. 243-270.

[9] S. CABAY AND P. KOSSOWSKI, *Power series remainder sequences and Padé fractions over integral domains*, J. Symbolic Computation, to appear.

[10] J. DURBIN, *The fitting of time-series models*, Rev. Inst. Internat. Statist., 28 (1960), pp. 233-244.

[11] O. I. ELGERD, *Control Systems Theory*, McGraw-Hill, New York, 1967.

[12] W. B. GRAGG, *The Padé table and its relation to certain algorithms of numerical analysis*, SIAM Rev., 14 (1972), pp. 1-61.

[13] G. LABAHN, *Matrix Padé approximants*, Master's thesis, Dept. of Computing Science, University of Alberta, Edmonton, Canada, 1986.

[14] N. LEVINSON, *The Wiener* RMS (*root mean square*) *error in filter design*, J. Math. and Phys., 25 (1947), pp. 261-278.

[15] G. NEMETH AND M. ZIMANYI, *Polynomial type Padé approximants*, Math. Comp., 38 (1982), pp. 553-565.

[16] J. RISSANEN, *Solution of linear equations with Hankel and Toeplitz matrices*, Numer. Math., 22 (1974), pp. 361-366.

[17] ———, *Recursive evaluation of Padé approximants for matrix sequences*, IBM J. Res. Develop. (1972), pp. 401-406.

[18] ———, *Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrix polynomials*, Math. Comp., 27 (1973), pp. 147-154.

[19] Y. STARKAND, *Explicit formulas for matrix-valued Padé approximants*, J. Comp. and Appl. Math., 5 (1979), pp. 63-65.

[20] W. F. TRENCH, *An algorithm for the inversion of finite Hankel matrices*, SIAM J. Appl. Math., 13 (1965), pp. 1102-1107.

[21] P. WYNN, *The rational approximation of functions which are formally defined by a power series expansion*, Math. Comp., 14 (1960), pp. 147-186.