50 copies
89-09
George March 1989
Labahn's
grant #1
Math covers

Printing Requisition/Graphic Services

60806

1. Please complete unshaded areas on form as applicable.
2. Distribute copies as follows: White and Yellow to Graphic Services. Retain Pink Copies for your records.
3. On completion of order the Yellow copy will be returned with the printed material.
4. Please direct enquiries, quoting requisition number and account number, to extension 3451.

TITLE OR DESCRIPTION: A Fast, Reliable Algorithm for calculating Pade-Hermite Forms CS-89-09

DATE REQUISITIONED: March 17/89
DATE REQUIRED: ASAP
ACCOUNT NO.: 1 2 9 6 8 0 2 4 1

REQUISITIONER - PRINT: G. Labahn
PHONE: 4667
SIGNING AUTHORITY: S. DeAngelis / G. Labahn

MAILING INFO – NAME: Sue DeAngelis
DEPT.: C.S.
BLDG. & ROOM NO.: DC 2314
[X] DELIVER [ ] PICK-UP

Copyright: I hereby agree to assume all responsibility and liability for any infringement of copyrights and/or patent rights which may arise from the processing of, and reproduction of, any of the materials herein requested. I further agree to indemnify and hold blameless the University of Waterloo from any liability which may arise from said processing or reproducing. I also acknowledge that materials processed as a result of this requisition are for educational use only.

NUMBER OF PAGES: 40
NUMBER OF COPIES: 50

TYPE OF PAPER STOCK: [X] BOND [ ] NCR ___ PT. [X] COVER [ ] BRISTOL [X] SUPPLIED [ ]

PAPER SIZE: [X] 8½ x 11 [ ] 8½ x 14 [ ] 11 x 17 [ ]

PAPER COLOUR: [X] WHITE [ ]
INK: [X] BLACK [ ]

PRINTING: [ ] 1 SIDE ___ PGS. [X] 2 SIDES ___ PGS.
NUMBERING: FROM ___ TO ___

BINDING/FINISHING: 3 down left side
[X] COLLATING [X] STAPLING [ ] HOLE PUNCHED [ ] PLASTIC RING

FOLDING/PADDING
CUTTING SIZE

Special Instructions

Math fronts and backs enclosed.

COPY CENTRE — OPER. NO. / BLDG. / MACH. NO.

DESIGN & PASTE-UP — OPER. NO. / TIME / LABOUR CODE

OPER. NO.	TIME	LABOUR CODE
		D01
		D01
		D01

TYPESETTING — QUANTITY

	QUANTITY			
PAP00000				T01
PAP00000				T01
PAP00000				T01

PROOF

PRF				
PRF				
PRF				

NEGATIVES

	QUANTITY	OPER. NO.	TIME	LABOUR CODE
FLM				C01
FLM				C01
FLM				C01
FLM				C01
FLM				C01

PMT

PMT				C01
PMT				C01
PMT				C01

PLATES

PLT				P01
PLT				P01
PLT				P01

STOCK

				001
				001
				001
				001

BINDERY

RNG				B01
RNG				B01
RNG				B01
MIS00000				B01

OUTSIDE SERVICES

$ ___ COST

TAXES – PROVINCIAL [ ] FEDERAL [ ]

GRAPHIC SERV. OCT. 85 482-2

A Fast, Reliable Algorithm for calculating Padé-Hermite Forms

Stan Cabay
Dept. of Computing Science
University of Alberta

George Labahn
Department of Computer Science
University of Waterloo
Research Report CS-89-09

March, 1989

A Fast, Reliable Algorithm for calculating Padé-Hermite Forms

Stan Cabay †

Department of Computing Science
University of Alberta
Edmonton, Alberta
Canada
T6G 2H1

George Labahn

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

ABSTRACT

For a vector of k power series the notion of a Padé-Hermite residual sequence and its corresponding cofactor sequence are introduced and developed. Two new types of rational approximants, the weak Padé-Hermite form and the weak Padé-Hermite fraction, are introduced along with an associated residual and cofactor sequence. A recurrence relation and a subsequent piecewise offdiagonal algorithm for constructing the two cofactor and two residual sequences is presented.

The algorithm for constructing the two cofactor sequences results in a fast algorithm for calculating a Padé-Hermite approximant of any given type. When the vector of power series is normal, the algorithm is shown to calculate a Padé-Hermite form of type $(n_0, \cdots, n_k)$ in $O(k \cdot (n_0^2 + \cdots + n_k^2))$ operations. This complexity is the same as that of other fast algorithms for computing Padé-Hermite approximants. However, unlike other algorithms, the new algorithm also succeeds in the non-normal case, usually with only a moderate increase in cost.

Key words: Vector of power series, Padé-Hermite fraction, Padé-Hermite approximation rational approximation, Sylvester matrix

† Supported in part by NSERC grant #A8035

1. Introduction

Given a formal power series

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

with coefficients from a field F, a Padé approximant of type (m,n) for A(z) is a pair of polynomials (U(z),V(z)) of degrees at most m and n, respectively, satisfying

$$A(z)V(z) = U(z) + O(z^{m+n+1}). \tag{1.2}$$

We can think of (1.2) as

$$A(z) \approx \frac{U(z)}{V(z)}, \tag{1.3}$$

so in a sense a Padé approximant is a realization of the formal power series as a rational expression $\frac{U(z)}{V(z)}$, at least to a specific set of terms.

The notion of a Padé-Hermite approximation is somewhat similar. First, following Padé [12] in his classic thesis, we wish to select k+1 polynomials so that for y(z) as the given power series we have

$$P_0(z)y(z)^k + P_1(z)y(z)^{k-1} + \cdots + P_{k-1}(z)y(z) = P_k(z) + O(z^{n_0+\cdots+n_k+k}), \tag{1.4}$$

where the $\{n_i\}$ are the degrees of the polynomials $\{P_i(z)\}$. For example, when k=2, then (1.4) can be thought of as

$$P_0(z)\cdot y(z)^2 + P_1(z)\cdot y(z) - P_2(z) \approx 0 \tag{1.5}$$

that is,

$$y(z) \approx \frac{-P_1(z) + \sqrt{P_1(z)^2 + 4\cdot P_0(z)\cdot P_2(z)}}{2\cdot P_0(z)} \tag{1.6}$$

which is also a representation of y(z) as a rational expression. We could also equally well consider

$$P_0(z)\frac{d^k y(z)}{dz} + P_1(z)\frac{d^{k-1}y(z)}{dz} + \cdots + P_{k-1}(z)y(z) = P_k(z) + O(z^{n_0+\cdots+n_k+k}), \tag{1.7}$$

in which case we wish to realize y(z) as a power series solution of a linear differential equation, again up to at least a specific number of terms. Generalizing further, consider

$$P_0(z)A_0(z) + \cdots + A_k(z)P_k(z) = O(z^{n_0+ \cdots +n_k+k}), \tag{1.8}$$

where the $A_i(z)$ are any desired set of functions of the given formal power series of y(z) (it is usually true that we can further assume that $A_j(0) \neq 0$ for at least one value j). In this last example, the polynomials $(P_0(z), \cdots, P_k(z))$ define a Padé-Hermite approximant of type $(n_0, \cdots, n_k)$ for the given system of power series $(A_0(z), \cdots, A_k(z))$.

Padé-Hermite approximants were introduced by Della Dora and Dicrescenzo [4] as a generalization of the quadratic approximants of Shafer [14] and the D-Log approximants of Baker [1]. Both of these concepts, in turn, began with ideas that originated from the thesis of Padé and some previous work of Hermite[7].

In addition to introducing the concept of Padé-Hermite approximants, Della Dora and Dicrescenzo also defined the notion of a Padé-Hermite table. This is a generalization of the normal definition of the extended Padé table (c.f., Gragg [5]). Relationships between neighboring entries in the table were then discovered that provided an algorithm to calculate such approximants. Other relationships in the Padé-Hermite table, and subsequently an alternate algorithm to calculate these approximants, were also discovered by Paszkowski [13].

The resulting algorithms, however, cannot be applied to arbitrary power series. The algorithms of both Della Dora et al and Paszkowski require that the vector of power series be **normal** (c.f., Paszkowski [13]). Della Dora and Dicrescenzo call such vector power series **perfect** rather than normal. Related to the concept of a Padé-Hermite approximation is a linear system of equations having a generalized Sylvester matrix as its coefficients. The normality condition requires that the coefficient matrix, along with a specific set of submatrices, be nonsingular. The normality requirement is a strong one. For example, the constant terms of all the $A_i(z)$'s need be nonzero for the system to satisfy the normality condition.

In this paper, we present an algorithm to calculate a Padé-Hermite approximant of a given type. This algorithm can be applied to any vector of power series; no requirement of normality is needed. A new type of rational approximant, the weak Padé-Hermite approximant, introduced in this paper, is central to the success of this procedure. These are a type of multidimensional rational approximant that can be transformed, if so desired, into a set of simultaneous Padé approximants (c.f., de Bruin[2]) for the given set of power series. Also introduced in this paper is the concept of a normal point in the Padé-Hermite table. The calculation of the desired approximant is obtained by iterating from one normal point to the next along a piecewise linear path in the Padé-Hermite table. When $k = 1$, Padé-Hermite approximants reduce to Padé approximants, and the algorithm becomes that of Cabay and Choi [3] and the scalar algorithm of Labahn and Cabay[9]. When $k = 1$, and the input power series are polynomials, our iteration scheme has close ties with the Extended Euclidean Algorithm. Indeed, by reversing the order of the coefficients of the input polynomials and travelling along a specific path our algorithm reduces to the EEA for these polynomials.

A cost analysis is also provided, showing that the algorithm generally reduces the cost by one order of magnitude to other methods that succeed in the non-normal case. In the normal case, the algorithm is of the same complexity as the algorithms of Della Dora et al and Paszkowski.

2. Basic Definitions

For a given integer $k \geq 0$, let

$$A_i(z) = \sum_{j=0}^{\infty} a_{i,j}\, z^j \quad , \quad i = 0, \cdots, k, \tag{2.1}$$

be a set of k+1 formal power series with coefficients $a_{i,j}$ coming from a field F. For a vector of non-negative integers $(n_0, n_1, \cdots, n_k)$, let

$$P_i(z) = \sum_{j=0}^{n_i} p_{i,j}\, z^j \quad , \quad i = 0, \cdots, k, \tag{2.2}$$

be a set of k+1 polynomials.

Definition 2.1 (Della Dora and Dicrescenzo [4]): The vector of polynomials $(P_0(z), \cdots, P_k(z))$ is defined to be a **Padé-Hermite form (PHFo)** of type $(n_0, \cdots, n_k)$ for the vector of power series $(A_0(z), \cdots, A_k(z))$ if

I) $\partial(P_i(z)) \leq n_i$, for i = 0, ... , k, II)

$$\sum_{i=0}^{k} A_i(z)P_i(z) = z^{n_0+\cdots+n_k+k} \cdot R(z), \tag{2.3}$$

where R(z) is a power series, and

III) the $P_i(z)$ are not all identically 0.

■

R(z) is called the residual of type $(n_0, \cdots, n_k)$ for the vector of power series. When k=1, and $A_1(z) = -1$, then Definition 2.1 corresponds to the definition of a Padé form for the power series $A(z)=A_0(z)$ (c.f., Gragg [5]). When k=1, $A_0(z) = A'(z)$ and $A_1(z) = A(z)$, we obtain the D-Log approximant of Baker [1]. When k=2, $A_0(z) = A^2(z)$, $A_1(z) = A(z)$, and $A_2(z) = 1$, we obtain the quadratic approximation of Shafer [14].

We extend Definition 2.1 to allow n_i to take on the value -1, but where at least one n_j must still be nonnegative. When $n_i = -1$, we define $P_i(z) = 0$. This is equivalent to $A_i(z)$ being absent, i.e., we are determining a PHFo for k, rather than for k+1, power series. Thus, for example, solving

$$A^2(z)P(z) + Q(z) = O(z^{m+n+1}), \tag{2.4}$$

where P(z) and Q(z) are to have degrees at most m and n, respectively, is the same as determining Shafer's quadratic approximation of type (m,-1,n).

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

$$P(z) = p_0 + p_1 z + \cdots + p_n z^n, \tag{2.5}$$

we write P (i.e., the same symbol but without the z variable) to mean the n+1 by 1 vector

$$P = \left[p_0, \cdots, p_n\right]^t. \tag{2.6}$$

Let

$$S_{(n_0, \cdots, n_k)} = \begin{bmatrix} a_{0,0} & & & \vert & & \vert & a_{k,0} & & \\ & \cdot & & \vert & & \vert & & \cdot & \\ \cdot & & a_{0,0} & \vert & \cdots & \vert & \cdot & & a_{k,0} \\ \cdot & & \cdot & \vert & & \vert & \cdot & & \cdot \\ & \cdot & & \vert & & \vert & & \cdot & \cdot \\ a_{0,\lambda} & \cdot & a_{0,\lambda-n_0} & \vert & & \vert & a_{k,\lambda} & \cdot & a_{k,\lambda-n_k} \end{bmatrix}, \tag{2.7}$$

where

$$\lambda = n_0 + \cdots + n_k + k - 1, \tag{2.8}$$

denote a generalized Sylvester matrix of type $(n_0, \cdots, n_k)$ for the vector $(A_0(z), \cdots, A_k(z))$. By equating coefficients of z^p for $0 \le p \le \lambda$, equation (2.3) can be written as

$$S_{(n_0, \cdots, n_k)} \begin{bmatrix} P_0 \\ \cdot \\ \cdot \\ P_k \end{bmatrix} = 0. \tag{2.9}$$

Theorem 2.2 (Existence of Padé-Hermite forms): For any vector of power series $(A_0(z), \cdots, A_k(z))$ and for any vector of integers $(n_0, \cdots, n_k)$, where $n_i \ge -1$ for all i and $n_i \ge 0$ for at least one i, there exists a PHFo of type $(n_0, \cdots, n_k)$.

Proof: The homogeneous system of linear equations (2.9) has $(\lambda+1)$ equations and $(\lambda+2)$ unknowns, where $\lambda \ge -1$, and hence has a nontrivial solution. Let $[P_0, \cdots, P_k]^t$ be one such nontrivial solution vector, with P_i having n_i+1 components. Then $P_0(z), \cdots, P_k(z)$ determined according to the convention (2.5) and (2.6) satisfy (2.3). Clearly conditions I and III of Definition 2.1 are also satisfied, and hence (2.9) determines a PHFo of the correct type.

■

From the proof of Theorem 2.2, it follows that there are many possible choices for a PHFo of a given type. As in the case of Padé approximants, it is desirable to have PHFo's that are unique, at least up to a multiplicative constant.

With λ defined by (2.8), let

$$T_{(n_0,\cdots,n_k)} = \begin{bmatrix} a_{0,0} & & & \Big| & & \Big| & a_{k,0} & & \\ & \cdot & & \Big| & & \Big| & & \cdot & \\ & & a_{0,0} & \Big| & \cdots & \Big| & & & a_{k,0} \\ \cdot & & \cdot & & & & \cdot & & \cdot \\ \cdot & & \cdot & \Big| & & \Big| & \cdot & & \cdot \\ & \cdot & & \Big| & & \Big| & & \cdot & \\ a_{0,\lambda+1} & \cdot & a_{0,\lambda-n_0+1} & \Big| & & \Big| & a_{k,\lambda+1} & \cdot & a_{k,\lambda-n_k+1} \end{bmatrix}, \tag{2.10}$$

and set

$$d_{(n_0,\cdots,n_k)} = \begin{cases} 0, & \text{if } n_i = -1 \text{ for } all\ i, \\ det(T_{(n_0,\cdots,n_k)}), & otherwise. \end{cases} \tag{2.11}$$

Thus $T_{(n_0,\cdots,n_k)}$ is a $(\lambda+2) \times (\lambda+2)$ matrix formed by bordering $S_{(n_0,\ \cdots\ ,\ n_k)}$ with one additional row. Then, a sufficient condition for PHFo's to be unique up to a multiplicative constant is given by

Theorem 2.3. If $d_{(n_0,\cdots,n_k)} \neq 0$, then PHFo's are unique up to multiplication by a non-zero element of the field F. In addition, the first term, R(0), of the residual in condition II of a PHFo,

$$A_0(z)P_0(z) + \cdots + A_k(z)P_k(z) = z^{\lambda+1}R(z), \tag{2.12}$$

is nonzero.

Proof: Equation (2.3), (2.9) and (2.12) are each equivalent to

$$T_{(n_0,\cdots,n_k)} \begin{bmatrix} P_0 \\ \cdot \\ \cdot \\ P_k \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ r_0 \end{bmatrix}, \tag{2.13}$$

where $r_0 = R(0)$. Now suppose $r_0 = 0$. Then, because $T_{(n_0,\cdots,n_k)}$ is nonsingular, $P_i = 0$ for all $0 \leq i \leq k$. But then $P_i(z) = 0$ for all $0 \leq i \leq k$ which contradicts condition III of Definition 2.1. Hence, we must have that $r_0 \neq 0$ when $d_{(n_0,\cdots,n_k)} \neq 0$.

To show uniqueness up to a nonzero constant, suppose $(P_0(z), \cdots, P_k(z))$ and $(P_0'(z), \cdots, P_k'(z))$ are both PHFo's of the correct type. From the previous paragraph, the leading terms of the residuals, r_0 and r'_0 must be nonzero. Hence there exists a nonzero constant, c such that $c \cdot r_0 = r'_0$. Then, from (2.12), the vector

$$c \cdot \begin{bmatrix} P_0 \\ \cdot \\ \cdot \\ P_k \end{bmatrix} - \begin{bmatrix} P_0' \\ \cdot \\ \cdot \\ P_k' \end{bmatrix} \tag{2.14}$$

is a solution to a homogeneous system of equations with coefficient matrix $T_{(n_0, \cdots, n_k)}$. Since by assumption this matrix is nonsingular, the solution must be zero. From this it is easy to see that

$$c \cdot P_i(z) = P_i'(z) \tag{2.15}$$

for all i. Hence the PHFo is unique up to multiplication by a nonzero element of the field F.

■

Weak Padé-Hermite Forms and **Fractions**

For the rest of this paper we will assume that $a_{i,0} \neq 0$ for at least one i (this must be true, for example, if ever $d_{(n_0, \cdots, n_k)} \neq 0$). For clarity of presentation and without a loss of generality, we assume this is true for i=0, i.e., $a_{0,0} \neq 0$. Set

$$A(z) = A_0(z), \ \ B(z) = (A_1(z), \cdots, A_k(z)) \tag{3.1}$$

where we view B(z) as a $1 \times k$ matrix of power series. Let P(z) be a polynomial and Q(z) be a $k \times 1$ matrix polynomial. Then, Definition 2.1 can be expressed alternately in matrix form as

Definition 3.1. The pair (P(z),Q(z)) is a **Padé-Hermite Form (PHFo)** for $(A(z),B(z))$ of type $(n_0, \cdots, n_k)$ if

I) $\partial(P(z)) \leq n_0$ and $\partial_{(j)}(Q(z)) \leq n_j$, where $\partial_{(j)}$ denotes the degree of the j th row,

II) $A(z) \cdot P(z) \ + \ B(z) \cdot Q(z) = z^{n_0 + \cdots + n_k + k} \cdot R(z),$ (3.2) where R(z) is a power series, and

III) at least one of P(z) or Q(z) is nonzero.

■

That Definition 3.1 and Definition 2.1 are equivalent follows by setting

$$P(z) = P_0(z)\ , \quad Q(z) = \begin{bmatrix} P_1(z) \\ \cdot \\ P_k(z) \end{bmatrix} \tag{3.3}$$

in Definition 3.1.

The fact that $a_{0,0} \neq 0$ has some implications, which will prove useful later. If $n_i = -1$, for $i = 1, \cdots, k$, then $n_0 \geq 0$ and the PHFo is given trivially by

$$(P_0(z), \cdots, P_k(z)) = (z^{n_0}, 0, , \cdots, 0). \tag{3.4}$$

On the other hand, if $n_i \geq 0$ for some i, $1 \leq i \leq k$, then (3.2) together with the assumption that $a_{0,0} \neq 0$ implies that $Q(z) \neq 0$. In this case, condition III in Definition 3.1 and 3.2 can be replaced by the simpler requirement that $Q(z) \neq 0$.

As noted in Section 2, PHFo's are equivalent to solutions of the linear system of equations (2.9) having $S_{(n_0, \cdots, n_k)}$ as a coefficient matrix. Consider now the linear system that results from the deletion of the last $k - 1$ rows of the Sylvester matrix (2.7). The system (2.9) now is guaranteed to have, not one, but at least k linearly independent solutions. Each solution still satisfies (3.2) but only with a relaxed order condition. Thus, it is only a kind of PHFo, defined in this weakener sense. Such solutions are introduced in this paper primarily to facilitate the development in Section 5 of an algorithm for computing the genuine PHFo's satisfying Definition 3.1. Arranging k such solutions by columns, formally, we have

Definition 3.2. Let $a_{0,0} \neq 0$ and let U(z) and V(z) be matrix polynomials of size $1 \times k$ and $k \times k$, respectively. The pair (U(z),V(z)) is a **Weak Padé-Hermite Form (WPHFo)** for $(A(z),B(z))$ of type $(n_0, \cdots, n_k)$, where $n_i \geq 0$ for $0 \leq i \leq k$, if

I) $\partial(U(z)) \leq n_0$ and $\partial_{(j)}(V(z)) \leq n_j$, where $\partial_{(j)}$ denotes degree of the j th row,

II) $A(z)\cdot U(z) + B(z)\cdot V(z) = z^{n_0 + \cdots + n_k + 1}\cdot W(z)$ (3.5)

where W(z) is a $1\times k$ matrix of power series, and

III) the columns of $V_i(z)$ are linearly independent.

■

The matrix polynomials U(z), V(z), and W(z) will be called the weak Padé-Hermite numerator, denominator, and residual (all of type $(n_0, \cdots, n_k)$), respectively. When k=1, Definition 3.2 is the scalar definition of a Padé form (c.f., Gragg [5]). Note that, just as for PHFo's, the notion of a WPHFo can be defined also for the case where $n_i = -1$ for some i. This extension is not required for the following development, however, so we will not pursue it further.

Let

$$V(z) = \begin{bmatrix} V_1(z) \\ \cdot \\ \cdot \\ V_k(z) \end{bmatrix} \tag{3.6}$$

and

$$U(z) = \sum_{j=0}^{n_0} u_j z^j \ , \ V_i(z) = \sum_{j=0}^{n_i} v_{i,j} z^j, \tag{3.7}$$

where $u_j(z)$ and $v_{i,j}$ are $1\times k$ vectors of coefficients from the field F. Equation (3.5) can then be written as

$$A_0(z)\cdot U(z) + A_1(z)\cdot V_1(z) + \cdots + A_k(z)\cdot V_k(z) = z^{n_0 + \cdots n_k+1}\cdot W(z), \tag{3.8}$$

and hence

$$S^*_{(n_0, \cdots, n_k)}\cdot \begin{bmatrix} U \\ V_1 \\ \cdot \\ \cdot \\ V_k \end{bmatrix} = 0, \tag{3.9}$$

where $S^*_{(n_0, \cdots, n_k)}$ is given by

$$S^*_{(n_0, \cdots, n_k)} = \left[\begin{array}{ccc|c|ccc} a_{0,0} & & & & a_{k,0} & & \\ \cdot & & & & \cdot & & a_{k,0} \\ \cdot & a_{0,0} & & \cdots & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & & \\ a_{0,\eta} & \cdot & a_{0,\eta-n_0} & & a_{k,\eta} & \cdot & a_{k,\eta-n_k} \end{array}\right] \tag{3.10}$$

and $\eta = n_0 + \cdots + n_k$.

Theorem 3.3 (Existence of WPHFo's): For any pair $(A(z),B(z))$ and $(n_0, \cdots, n_k)$, where $n_i \geq 0$, $0 \leq i \leq k$,, there exists a WPHFo of type $(n_0, \cdots, n_k)$.

Proof: The proof of Theorem 3.3 parallels the proof of Theorem 2.2. The system of equations (3.9) having $S^*_{(n_0, \cdots, n_k)}$ as its coefficient matrix has $\eta+1$ equations and $\eta+k+1$ unknowns. Thus (3.9) must have at least k linearly independent solutions (of course there may be more). Any set of k linearly independent vectors can be arranged by columns to yield a solution $[U^t, V_1^t, \cdots, V_k^t]^t$ of (3.9).

The contention is that the columns of $[V_1^t, \cdots, V_k^t]^t$ are also linearly independent. For suppose that there exists a nontrivial k-tuple, $\alpha = [\alpha_1, \cdots, \alpha_k]^t$, where α_i are from the field F, such that

$$[V_1^t, \cdots, V_k^t]^t \alpha = 0. \tag{3.11}$$

Then, multiplying (3.9) on both sides by α yields

$$\begin{bmatrix} a_{0,0} & & \\ \cdot & \ddots & \\ \cdot & & \\ a_{0,n_0} & & a_{0,0} \\ \cdot & & \cdot \\ a_{0,\eta} & & a_{0,\eta-n_0} \end{bmatrix} \cdot U \cdot \alpha = 0. \tag{3.12}$$

Thus,

$$U \cdot \alpha = 0 \tag{3.13}$$

since $a_{0,0} \neq 0$. Equations (3.11) and (3.13) then together contradict the assumption that the columns of $[U^t, V_1^t, \cdots, V_k^t]^t$ are linearly independent.

Organizing the matrix polynomials U(z) and V(z) according to (3.6) and (3.7) and using the conventions of (2.5) and (2.6) determines a WPHFo of the correct type.

■

From the proof of Theorem 3.3, it follows that if $S^*_{(n_0, \cdots, n_k)}$ has maximal rank, then WPHFo's are unique up to multiplication of U(z) and V(z) on the right by a nonsingular $k \times k$ matrix. On the other hand, if the rank of $S^*_{(n_0, \cdots, n_k)}$ is less than maximal, then more than one independent WPHFo exists.

Definition 3.4: A pair (U(z),V(z)) is a **weak Padé-Hermite fraction (WPHFr)** for (A(z),B(z)) of type $(n_0, \cdots, n_k)$ if

I) it is a WPHFo of type $(n_0, \cdots, n_k)$ for $(A(z),B(z))$, and

II) the constant term, V(0), of the weak Padé-Hermite denominator is a nonsingular matrix.

■

When $k = 1$, a WPHFr is the same as a scalar Padé fraction (c.f., Gragg[5]). Note that, unlike the case for WPHFo's, a WPHFr can only be defined when all the n_i are nonnegative integers.

A WPHFr can be interpreted as providing a a set of simultaneous Padé approximants for the quotient power series $\frac{A_i(z)}{A_0(z)}$ (c.f., de Bruin[2]). Indeed, since V(0) is nonsingular, the inverse of the matrix polynomial $V(z)$ can be determined as a matrix power series. Thus, we obtain

$$\frac{B(z)}{A_0(z)} \approx -U(z) \cdot V(z)^{-1}. \tag{3.14}$$

Since

$$U(z) \cdot V(z)^{-1} = U(z) \cdot adj(V(z))/det(V(z)), \tag{3.15}$$

equations (3.14) and (3.15) give a simultaneous rational approximation for each power series

$$\frac{A_i(z)}{A_0(z)} \approx \frac{N_i(z)}{D(z)}, \; i = 1, .., k. \tag{3.16}$$

It is not difficult to see that $N_i(z)$ has at most degree $N - n_i$ and that D(z) has at most degree $N - n_0$, where $N = n_0 + \cdots + n_k$. Hence, the polynomials $(D(z), N_1(z), \cdots, N_k(z))$ form a set of simultaneous Padé approximants to the power series $A_1(z)/A_0(z), \cdots, A_k(z)/A_0(z)$ of type $(n_0, \cdots, n_k)$. This can also be represented as a solution to the German polynomial approximation problem of type $(n_0, \cdots, n_k)$ for the power series $(A_0(z), \cdots, A_k(z))$ or as directed vector Padé approximants for the vector of power series $(A_0(z), \cdots, A_k(z))$ in the unit coordinate directions (c.f., Graves-Morris[6]).

Finally, it can also be shown that the approximation (3.14) is also a reduced approximation; that is, there are no nontrivial common right divisors of both U(z) and V(z) (a nontrivial matrix polynomial is one whose determinant degree is greater than 0).

Unlike WPHFo's, a WPHFr of a given type need not always exist. However, in a manner similar to that of PHFo's, we obtain a sufficient condition for existence of a WPHFr in certain cases.

Let

$$T^*_{(n_0, \cdots, n_k)} = \left[\begin{array}{ccc|cc|c|cc} a_{0,0} & 0 & & 0 & & & 0 & \\ . & a_{0,0} & & a_{1,0} & & & a_{k,0} & \\ . & . & & . & & & . & \\ . & . & & . & a_{1,0} & \cdots & . & a_{k,0} \\ . & . & a_{0,0} & . & . & & . & . \\ . & . & . & . & . & & . & . \\ & & & & & & & \\ a_{0,\eta} & a_{0,\eta-1} & a_{0,\eta-n_0} & a_{1,\eta-1} & a_{1,\eta-n_1} & & a_{k,\eta-1} & a_{k,\eta-n_k} \end{array} \right], \tag{3.17}$$

and set

$$d^*_{(n_0, \cdots, n_k)} = det(T^*_{(n_0, \cdots, n_k)}). \tag{3.18}$$

We then have

Theorem 3.5. If $d^*_{(n_0 n_1, \cdots, n_k)} \neq 0$, then every WPHFo of type $(n_0, \cdots, n_k)$ is a WPHFr. In addition, any WPHFr of type $(n_0, \cdots, n_k)$ is unique up to multiplication on the right by a nonsingular $k \times k$ matrix with coefficients from the field F.

Proof: Equation (3.9) may be rewritten as

$$T^{*}_{(n_0,\,\cdots,\,n_k)}\cdot\begin{bmatrix} u_0 \\ \cdot \\ u_{n_0} \\ v_{1,1} \\ \\ v_{1,n_1} \\ \cdot \\ v_{k,1} \\ \cdot \\ v_{k,n_k}\end{bmatrix} = -\begin{bmatrix} a_{1,0} & & a_{k,0} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{1,\eta} & & a_{k,\eta}\end{bmatrix}\cdot\begin{bmatrix} v_{1,0} \\ \cdot \\ \cdot \\ v_{k,0}\end{bmatrix}. \tag{3.19}$$

Therefore $d^{*}_{(n_0,\cdots,n_k)} \neq 0$ implies that all the solutions of (3.19) can be obtained by first assigning arbitrarily the $k \times k$ matrix

$$V(0) = \begin{bmatrix} v_{1,0} \\ \cdot \\ \cdot \\ v_{k,0}\end{bmatrix} \tag{3.20}$$

(recall that each $v_{i,j}$ represents a $1 \times k$ vector) and then solving for the remaining components $u_0, \cdots, u_{n_k}, v_{1,1}, \cdots, v_{k,n_k}$. If V(0) is chosen to be a singular matrix, then the solution obtained by solving (3.19) violates condition III in the definition of a WPHFo. Thus, in this case, all WPHFo's are in fact WPHFr's.

Note that equation (3.19) still holds in the special case when $n_i = 0$ for any i, $1 \leq i \leq k$, even though there are no $v_{i,j}$ in the left side of (3.19) in this case. For such an i, there are also no $a_{i,j}$ in $T^{*}_{(n_0,\cdots,n_k)}$, yet the right side of (3.19) can still be formed. The rest of the arguments above also hold in this special case.

To show uniqueness, suppose (U(z),V(z)) and $(U^{*}(z),V^{*}(z))$ are two WPHFr's for $(A(z),B(z))$. Then, the corresponding matrices V(0) and $V'(0)$ in (3.20) for each solution are both nonsingular matrices with coefficients from the field F. Thus, there exists a nonsingular matrix, C, with coefficients from F, satisfying

$$\begin{bmatrix} v_{1,0} \\ \cdot \\ \cdot \\ v_{k,0} \end{bmatrix} = \begin{bmatrix} v^*_{1,0} \\ \cdot \\ \cdot \\ v^*_{k,0} \end{bmatrix} \cdot C. \tag{3.21}$$

It follows from (3.19) that

$$V(z) = V^*(z) \cdot C \quad \text{and} \quad U(z) = U^*(z) \cdot C, \tag{3.22}$$

and so uniqueness holds.

■

Up to this point Sections 2 and 3 have only their similarity in development in common. However, for a vector of integers $(n_0, \cdots, n_k)$ where $n_i \geq 0$, $1 \leq i \leq k$, we have

$$T_{(n_0-1, \cdots, n_k-1)} = \left[\begin{array}{cc|cc|c|cc} a_{0,0} & & a_{1,0} & & & a_{k,0} & \\ \cdot & & \cdot & & & \cdot & \\ \cdot & & \cdot & a_{1,0} & & \cdot & a_{k,0} \\ \cdot & a_{0,0} & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & & & & & & \\ a_{0,\eta-1} & a_{0,\eta-n_0} & a_{1,\eta-1} & a_{1,\eta-n_1} & & a_{k,\eta-1} & a_{k,\eta-n_k} \end{array}\right] \tag{3.23}$$

which is the same as $T^*_{(n_0, \cdots, n_k)}$ with the first row and column eliminated. Thus,

$$d^*_{(n_0, \cdots, n_k)} = \begin{cases} a_{0,0}, & \text{if } n_i = 0 \text{ for } \textit{all } i, \\ a_{0,0} \cdot d_{(n_0-1, \cdots, n_k-1)}, & \textit{otherwise}. \end{cases} \tag{3.24}$$

Central to our results is the fact that Theorems 2.3 and 3.5 actually classifies when the Sylvester determinant is nonzero. Indeed we have

Theorem 3.6. Let (A(z),B(z)) be as in (3.1) and let $(n_0, \cdots, n_k)$ with $n_i \geq -1$, be a vector of integers. Then $d_{(n_0, \cdots, n_k)} \neq 0$ if and only if

1) there exists a PHFo (P(z),Q(z)) of type $(n_0, \cdots, n_k)$ having a nonzero leading term in the residual, and

2) there exists a WPHFr of type $(n_0+1, \cdots, n_k+1)$ for (A(z),B(z)).

Proof: We show that if 1) and 2) are true, then $d_{(n_0, \cdots, n_k)} \neq 0$. The converse is clearly true from Theorems 2.3 and 3.5. Suppose there are solutions to equations (2.13) and (3.19) (with $(n_0, \cdots, n_k)$ replaced by $(n_0+1, \cdots, n_k+1)$), but that

$$d_{(n_0-1, \cdots, n_k-1)} = 0, \tag{3.27}$$

that is, suppose $T_{(n_0, \cdots, n_k)}$ is singular. Then, there exists a nontrivial solution to the homogeneous system of equations

$$(s_0, \cdots, s_{\lambda+1}) \cdot T_{(n_0, \cdots, n_k)} = 0. \tag{3.28}$$

Equation (2.13) together with (3.28) then yield

$$s_{\lambda+1} \cdot r_0 = (s_0, \cdots, s_{\lambda+1}) \cdot T_{(n_0, \cdots, n_k)} \cdot \begin{bmatrix} p_{0,0} \\ \cdot \\ \cdot \\ \cdot \\ p_{k,n_k} \end{bmatrix} = 0. \tag{3.29}$$

Hence $s_{\lambda+1} = 0$, since by assumption (1), $r_0 \neq 0$.

Given that $a_{0,0} \neq 0$, let us determine τ such that

$$(\tau, s_0, \cdots, s_\lambda+1) \cdot (a_{0,0}, \cdots, a_{0,\lambda+2})^t = 0. \tag{3.30}$$

Then,

$$(\tau, s_0, \cdots, s_\lambda+1) \cdot T^*_{(n_0+1, \cdots, n_k+1)} = 0, \tag{3.31}$$

where $T^*_{(n_0+1, \cdots, n_k+1)}$ is given by (3.17). Note that with $(n_0, \cdots, n_k)$ replaced by $(n_0+1, \cdots, n_k+1)$, $\eta = \lambda + 2$ in (3.17). Equations (3.19) and (3.31) then yield

$$(\tau,\, s_0,\, \cdots, s_{\lambda+1})\cdot\begin{bmatrix} a_{1,0} & & a_{k,0} \\ \cdot & & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & & \cdot \\ a_{1,\lambda+2} & & a_{k,\lambda+2} \end{bmatrix}\cdot\begin{bmatrix} v_{1,0} \\ \cdot \\ \cdot \\ v_{k,0} \end{bmatrix} = 0, \tag{3.32}$$

where, by assumption (2),

$$V(0) = \begin{bmatrix} v_{1,0} \\ \cdot \\ v_{k,0} \end{bmatrix} \tag{3.33}$$

is a nonsingular matrix. Hence,

$$(\tau,\, s_0,\, \cdots, s_{\lambda+1})\cdot\begin{bmatrix} a_{1,0} & & a_{k,0} \\ \cdot & & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & & \cdot \\ a_{1,\lambda+2} & & a_{k,\lambda+2} \end{bmatrix} = 0. \tag{3.34}$$

Therefore, from (3.28), (3.30) and (3.34), we get

$$(\tau,\, s_0,\, \cdots, s_{\lambda+1})\cdot\left[\begin{array}{cc|c|cc} a_{0,0} & & & a_{k,0} & \\ \cdot & & & \cdot & \\ \cdot & & & \cdot & a_{k,0} \\ \cdot & a_{0,0} & \cdots & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ & & & & \\ a_{0,\lambda+2} & a_{0,\lambda-n_0+2)} & & a_{k,\lambda+2} & a_{k,\lambda-n_k+2} \end{array}\right] = 0. \tag{3.35}$$

Equation (3.35) coupled with $s_{\lambda+1} = 0$ implies that

$$(\tau,\, s_0,\, \cdots, s_\lambda)\cdot T_{(n_0, \cdots, n_k)} = 0. \tag{3.36}$$

An induction argument can then be used to show that

$$s_{\lambda+1} = s_\lambda = \cdots s_0 = 0, \tag{3.37}$$

which contradicts the assumption that there is a nontrivial solution to (3.28). Hence, $T_{(n_0, \cdots, n_k)}$ is nonsingular.

The converse of Theorem 3.6 is clear from Theorems 2.3 and 3.5.

■

Remark 1: When k = 1, Theorem 3.6 was proved in [10].

Remark 2: Notice that the proof of Theorem 3.6 gives necessary and sufficient conditions for the nonsingularity of a generalized Sylvester matrix. Indeed, if

$$S = \begin{bmatrix} s_{0,0} & s_{0,m_0} & \Big| & & \Big| & s_{k,0} & s_{k,m_k} \\ \cdot & \cdot & \Big| & & \Big| & \cdot & \cdot \\ \cdot & \cdot & \Big| & \cdots & \Big| & \cdot & \cdot \\ \cdot & \cdot & \Big| & & \Big| & \cdot & \cdot \\ s_{0,N} & s_{0,N+m_0} & & & & s_{k,N} & s_{k,N+m_k} \end{bmatrix} \tag{3.38}$$

with $N = m_0 + \cdots + m_k + k$, then, using arguments similar to those used in the proof of Theorem 3.6, it can be seen that S is nonsingular if and only if there exist solutions to the equations

$$S \cdot (x_0^{(i)}, \cdots, x_N^{(i)})^t = -(s_{i,m_i+1}, \cdots, s_{i,N+m_i+1})^t \quad \text{for} \quad i \in 0, \cdots, k \tag{3.39}$$

$$S \cdot (y_0, \cdots, y_N)^t = (0, \cdots, 0, 1)^t \tag{3.40}$$

4. Padé-Hermite and Weak Padé-Hermite Residual Sequences

Following Della Dora and Dicrescenzo, we define a **Padé-Hermite Table** for a $1 \times (k+1)$ vector of power series (A(z),B(z)) to be an infinite (k+1)-dimensional collection of Padé-Hermite forms of type $(n_0, \cdots, n_k)$ for $n_i = -1, 0, 1, 2, \cdots$. It is assumed that, with the exception of the case where $(n_0, \cdots, n_k) = (-1, \cdots, -1)$, there is precisely one entry assigned to each position in the table. From Theorem 2.2, it follows that a Padé-Hermite table exists for any given $(A_0(z), \cdots, A_k(z))$. However the table is not unique. This is unlike the definition of the usual Padé table (when k = 1), since there a Padé table consists of a collection of Padé fractions (c.f.[9]), which are unique.

Corresponding to a similar notion introduced for the usual Padé approximants, a vector of power series $(A_0(z), \cdots, A_k(z))$ is said to be **normal** (c.f., Paszkowski[13]) if $d_{(n_0, \cdots, n_k)} \neq 0$ for all n_i. (Della Dora and Dicrescenzo [4] use the term **perfect** to describe this property). When the vector of power series is normal, it follows from Theorem 2.3 that every entry in the Padé-Hermite table is unique up to multiplicative constant. Also from Theorem 2.3 it can be seen that the Padé-Hermite table for normal vectors of power series can be made unique by insisting that the first term, R(0), of the residual be set to 1.

Notice that any vector of integers $(n_0, \cdots, n_k)$, where $n_i \geq 0$ for at least one i, can be associated with a point, $(m_0, \cdots, m_k)$ in the Padé-Hermite table by setting

$$m_j = \begin{cases} n_j & \text{if } n_j \geq -1, \\ -1 & \textit{otherwise.} \end{cases} \tag{4.1}$$

The point $(m_0, \cdots, m_k)$ is called the **representative** of $(n_0, \cdots, n_k)$ in the Padé-Hermite table. When $n_i < 0$, the representative point in the Padé-Hermite table excludes the power series $A_i(z)$. Correspondingly, we also extend the definition of the Sylvester determinant (2.11) to arbitrary vectors of integers by setting

$$d_{(n_0, \cdots, n_k)} = d_{(m_0, \cdots, m_k)}, \tag{4.2}$$

that is, the Sylvester determinant becomes that of its representative.

Given a vector of power series (2.1) and a vector of integers $(n_0, \cdots, n_k)$, a corresponding PHFo can be determined by solving (2.9) using Gaussian elimination, say. This has the advantage that there need be no restriction on the input vector of power series. A similar remark may be made about the calculation of WPHFo's via the solution to the system (3.9). However, such calculations do not take into account the special structure of the coefficient matrices of the systems. The goal of this section is to describe a recurrence relation that will lead to an efficient algorithm for both the determination of a PHFo or a WPHFo of any type. The resulting algorithm will take advantage of the special structure of the coefficient matrix of (2.7) and (3.10), and at the same time it will not require any restrictions on the input. In particular, the assumption of normality will not be required.

Given a vector of power series (2.1), along with a vector $(n_0, \cdots, n_k)$ of nonnegative integers, permute the components so that

$$A_0(0) \neq 0, \cdots, A_l(0) \neq 0 \text{ and } A_j(0) = 0, \text{ for } j > l \tag{4.3}$$

and

$$n_0 \geq \cdots \geq n_l, \text{ and } n_{l+1} \leq \cdots \leq n_k. \tag{4.4}$$

This ordering is for presentation purposes only. If $A_i(0) = 0$, $0 \leq i \leq k$, it is only necessary to remove the largest factor, z^β, from all the power series. Any PHFo or WPHFo of type $(n_0, \cdots, n_k)$ for $(z^{-\beta} \cdot A_0(z), \cdots, z^{-\beta} \cdot A_k(z))$ is then also a PHFo or WPHFo, respectively, of the same type for $(A_0(z), \cdots, A_k(z))$.

We introduce a sequence of integer vectors in k+1 space

$$(n_0^{(0)}, \cdots, n_k^{(0)}), \ (n_0^{(1)}, \cdots, n_k^{(1)}), \ (n_0^{(2)}, \cdots, n_k^{(2)}), \quad \cdots \tag{4.5}$$

by setting

$$(n_0^{(0)}, \cdots, n_k^{(0)}) = (n_0 - M, n_1 - M, \cdots, n_k - M) \tag{4.6}$$

and

$$(n_0^{(i+1)}, \cdots, n_k^{(i+1)}) = (n_0^{(i)}, \cdots, n_k^{(i)}) + (s_i, \cdots, s_i) \tag{4.7}$$

where $s_i \geq 1$ and

$$M = \begin{cases} n_0 + 1, & \text{if } n_k \geq n_0, \\ max(n_1, n_k) + 2, & otherwise. \end{cases} \tag{4.8}$$

In (4.7), the s_i, $i \geq 0$, are selected so that

$$d_{(n_0^{(i+1)}, \cdots, n_k^{(i+1)})} \neq 0 \tag{4.9}$$

and

$$d_{(n_0^{(i)}+j,\ \cdots\ ,n_k^{(i)}+j)} = 0,\ j = 1,\ \cdots\ ,\ s_i-1. \tag{4.10}$$

Observe that the ordering (4.4) implies that $n_0^{(1)} \geq 0$, since otherwise the Sylvester matrix would have at least one zero row. The requirement that $n_0^{(i)} \geq 0$, $i \geq 1$ is important in the recurrence relation described latter in this section. Also, note that

$$(n_1^{(i)} - n_0^{(i)},\ \cdots\ ,\ n_k^{(i)} - n_0^{(i)}) = (n_1-n_0,\ \cdots\ ,n_k-n_0), \tag{4.11}$$

for all i, and consequently the sequence (4.5) lies along a straight line in k+1 space. The points $(n_0^{(i)},\ \cdots\ ,n_k^{(i)})$, for $i \geq 1$, are called the **nonsingular points** along the off-diagonal line starting at $(n_0^{(0)},\ \cdots\ ,n_k^{(0)})$ in the direction $(M,\ \cdots\ ,M)$.

The sequence of nonsingular points (4.5) determines a sequence

$$(m_0^{(0)},\ \cdots\ ,\ m_k^{(0)}),\ (m_0^{(1)},\ \cdots\ ,\ m_k^{(1)}),\ (m_0^{(2)},\ \cdots\ ,\ m_k^{(2)}),\ \ \cdots \tag{4.12}$$

of points in the Padé-Hermite table, where $(m_0^{(i)},\ \cdots\ ,\ m_k^{(i)})$ is the representative of $(n_0^{(i)},\ \cdots\ ,n_k^{(i)})$. The resulting points are called **normal** points in the Padé-Hermite table. The normal points representing the sequence (4.5) lie on a piecewise linear path in the Padé-Hermite table of (A(z),B(z)).

Example 4.1: Let k=3 and let (n_0,n_1,n_2,n_3) = (7,4,1,9) (hence, according to the ordering (4.3), $A_3(0)$ must be 0). Then $M = 8$, $(n_0^{(0)},n_1^{(0)},n_2^{(0)},n_3^{(0)}) = (-1,-4,-7,1)$, and the nonsingular points $(n_0^{(i)},n_1^{(i)},n_2^{(i)},n_3^{(i)})$ $i \geq 1$, lie along the off-diagonal line starting at (0,-3,-6,2) in the direction (1,1,1,1), that is, along the straight line in 4-space from (0,-3,-6,2) to (7,4,1,9). The corresponding representative points $(m_0^{(i)},m_1^{(i)},m_2^{(i)},m_3^{(i)})$, $i \geq 1$, then lie along the piecewise linear path defined by the line segments from (0,-1,-1,2) to (2,-1,-1,4), from (2,-1,-1,4) to (5,2,-1,7), and from (5,2,-1,7) to (7,4,1,9).

■

For i = 1, 2, ..., let $(P^{(i)}(z),Q^{(i)}(z))$ be a PHFo of type $(m_0^{(i)},\ \cdots\ ,\ m_k^{(i)})$ for (A(z),B(z)). Thus, according to Theorem 2.3, there exists a power series $R^{(i)}(z)$ such that

$$A(z)P^{(i)}(z) + B(z)\cdot Q^{(i)}(z) = z^{m_0^{(i)} + \cdots m_k^{(i)} + k} R^{(i)}(z) \tag{4.13}$$

where $R^{(i)}(0) \neq 0$. This PHFo is made unique by insisting that $R^{(i)}(0) = 1$ (c.f. Theorem 2.3).

Definition 4.2. The sequence

$$\left\{R^{(i)}(z)\right\}, \quad i = 1, 2, \cdots, \tag{4.14}$$

with $R^{(i)}(0) = 1$ is called the **Padé-Hermite residual sequence** for the vector of power series (A(z),B(z)). The sequence

$$\left\{(P^{(i)}(z), Q^{(i)}(z))\right\}, \quad i = 1, 2, \cdots, \tag{4.15}$$

is called the **Padé-Hermite cofactor sequence.**

■

Similarly, for $i = 1, 2, \ldots$, let $(U^{(i)}(z), V^{(i)}(z))$ be a WPHFr of type $(m_0^{(i)}+1, \cdots, m_k^{(i)}+1)$ for (A(z),B(z)). Then, there exists a vector of power series $W^{(i)}(z)$ such that

$$A(z)U^{(i)}(z) + B(z)\cdot V^{(i)}(z) = z^{m_0^{(i)} + \cdots m_k^{(i)} + k + 2} W^{(i)}(z) \tag{4.16}$$

where $det(V^{(i)}(0)) \neq 0$. This WPHFr is made unique by insisting that $V^{(i)}(0) = I$ (c.f. Theorem 3.5).

Definition 4.3. The sequence

$$\left\{W^{(i)}(z)\right\}, \quad i = 1, 2, \cdots, \tag{4.17}$$

is called the **weak Padé-Hermite residual sequence** for the vector of power series (A(z),B(z)). The corresponding sequence

$$\left\{(U^{(i)}(z), V^{(i)}(z))\right\}, \quad i = 1, 2, \cdots, \tag{4.18}$$

with $V^{(i)}(0) = I$ is called the corresponding **weak Padé-Hermite cofactor sequence.**

■

Example 4.4. Suppose $(n_0, \cdots, n_k)$ is ordered as in (4.4) and that $n_0 > n_1 \geq n_k$. Then $M = n_1 + 2$ and $(n_0^{(0)}, \cdots, n_k^{(0)}) = (n_0 - n_1 - 2, -2, n_2 - n_1 - 2, \cdots, n_k - n_1 - 2)$. Furthermore $s_0 = 1$, since $a_{0,0} \neq 0$

implies that the matrix (2.9) for the representative $(n_0-n_1-1, -1, \cdots, -1)$ of the point $(n_0^{(0)}, \cdots, n_k^{(0)})$ is

$$T_{(n_0-n_1-1,-1,\cdots,-1)} = \begin{bmatrix} a_{0,0} & \\ \cdot & \\ a_{0,n_0-n_1-1} & a_{0,0} \end{bmatrix}, \tag{4.19}$$

which has nonzero determinant. To obtain the PHFo of type $(m_0^{(1)}, \cdots, m_k^{(1)}) = (n_0-n_1-1,-1, \cdots,-1)$, we solve

$$\begin{bmatrix} a_{0,0} & \\ & \\ a_{0,n_0-n_1-1} & a_{0,0} \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ \\ p_{n_0-n_1-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; \tag{4.20}$$

and, so in this case,

$$P^{(1)}(z) = a_{0,0}^{-1} \cdot z^{n_0-n_1-1}, \quad Q^{(1)}(z) = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix}. \tag{4.21}$$

To determine the WPHFr of type $(n_0-n_1,0, \cdots,0)$, we solve

$$\begin{bmatrix} a_{0,0} & \\ & \\ a_{0,n_0-n_1} & a_{0,0} \end{bmatrix} \cdot \begin{bmatrix} u_0 \\ \\ u_{n_0-n} \end{bmatrix} = - \begin{bmatrix} a_{1,0} & a_{k,0} \\ & \\ a_{1,n_0-n_1} & a_{k,n_0-n_1} \end{bmatrix}. \tag{4.22}$$

Thus, in this case, the WPHFr of type $(n_0-n_1,0, \cdots,0)$, is given by

$$U^{(1)}(z) = -A^{-1}(z) \cdot B(z) \bmod z^{n_0-n_1+1} \text{ and } V^{(1)}(z) = I. \tag{4.23}$$

■

Examples 4.1 and 4.4 illustrate the effect of the definition of M in (4.8). When $n_k \geq n_0$, as in Example 4.1, a nonsingular point cannot occur for coordinates on the off-diagonal path smaller than $(0, n_1-n_0, \cdots, n_k-n_0)$. In (4.8), therefore M is selected so that $n_0^{(1)} \geq 0$. When $n_k < n_0$, as in Example 4.4, M is selected so that $(n_0^{(1)}, \cdots, n_k^{(1)})$ is the smallest point which has at least two coordinates in common with its representative. The desired effect, which has consequences in the subsequent development, is that in both cases $n_0^{(1)} \geq 0$ and $(n_0^{(1)}, \cdots, n_k^{(1)})$ has at least two coordinates in common with $(m_0^{(1)}, \cdots, m_k^{(1)})$.

Theorem 4.5: The Padé-Hermite and weak Pade-Hermite residual sequences and cofactor sequences (4.14), (4.15), (4.17) and (4.18) exist uniquely.

Proof: The result is a direct consequence of Theorem 3.6.

■

The algorithm described in Section 5 for constructing a PHFo of type $(n_0, \cdots, n_k)$ for (A(z),B(z)) involves the computation of all terms in the Padé-Hermite and weak Padé-Hermite cofactor sequences up to the point $(n_0, \cdots, n_k)$. Theorem 4.7 and the proof of Theorem 4.6 below give a relationship of the (i+1)-st terms of the sequences with the i-th terms, providing an effective mechanism for computing the sequences.

For each integer i, let

$$(\Delta_0^{(i)}, \cdots, \Delta_k^{(i)}) = (0,\ m_1^{(i)} - n_1^{(i)}, \cdots, m_k^{(i)} - n_k^{(i)}) \tag{4.24}$$

be the difference between the nonsingular point and its representation as a normal point. Because of the ordering (4.3) we have that

$$\Delta_0^{(i)} \leq \Delta_1^{(i)} \leq \cdots \leq \Delta_l^{(i)} \quad \text{and} \quad \Delta_{l+1}^{(i)} \geq \cdots \geq \Delta_k^{(i)}. \tag{4.25}$$

The main result of this section is

Theorem 4.6. Let $R^i(z)$ and $W^{(i)}(z)$ given by (4.13) and (4.16), respectively, be the residuals at the nonsingular point $(n_0, \cdots, n_k)$. For any integers $s \geq 1$ and $i \geq 1$, $(n_0^{(i)}+s, \cdots, n_k^{(i)}+s)$ is a nonsingular point for (A(z),B(z)) if and only if $(s-2,\ s-1-\Delta_1^{(i)}, \cdots, s-1-\Delta_k^{(i)})$ is a nonsingular point for $(R^{(i)}(z), W^{(i)}(z))$.

Proof: We first examine the representatives of all relevant points. Because of the ordering (4.4) we can define integers a and d so that, for $i \geq 1$,

$$n_j^{(i)} \geq -1,\ \ j=0,\ \cdots,a \quad \text{and} \quad j=d,\ \cdots,k$$

$$n_j^{(i)} < -1,\ \ j=a+1,\ \cdots,d-1. \tag{4.26}$$

Using (4.7) and (4.8) note that $a \geq 0$ and that either $a \geq 1$ or $d \leq k$. The representative of $(n_0^{(i)},\ \cdots,n_k^{(i)})$ is then

$$(m_0^{(i)},\cdots,m_k^{(i)}) = (n_0^{(i)},\cdots,n_a^{(i)},-1,\cdots,-1,\ n_d^{(i)},\cdots,n_k^{(i)}). \tag{4.27}$$

For $s > 0$, let b and c be integers such that

$$n_j^{(i)} + s \geq 0,\ \ j=0,\ \cdots,b \quad \text{and} \quad j=c,\ \cdots,k$$

$$n_j^{(i)} + s < 0,\ \ j=b+1,\ \cdots,c-1. \tag{4.28}$$

Clearly, from (4.26), $a \leq b \leq c \leq d$. The representative of $(n_0^{(i)}+s,\ \cdots,n_k^{(i)}+s)$ is therefore

$$(n_0^{(i)}+s,\cdots,n_b^{(i)}+s,-1,\cdots,-1,\ n_c^{(i)}+s,\cdots,n_k^{(i)}+s). \tag{4.29}$$

Finally, from (4.24) and (4.27), observe that

$$s-1-\Delta_j^{(i)} = s-1+n_j^{(i)}-m_j^{(i)} = \begin{cases} s-1, & 1 \leq j \leq a, \\ n_j^{(i)}+s, & a+1 \leq j \leq d-1, \\ s-1, & d \leq j \leq k. \end{cases} \tag{4.30}$$

Thus, from (4.28) and (4.30), it follows that the representative of $(s-2,s-1-\Delta_1^{(i)},\ \cdots,s-1-\Delta_k^{(i)})$ is

$$(s-2,s-1,\cdots,s-1,n_{a+1}^{(i)}+s,\cdots,n_b^{(i)}+s,-1,\cdots,-1,n_c^{(i)}+s,\cdots,n_{d-1}^{(i)}+s,s-1,\cdots,s-1) \tag{4.31}$$

By assumption, $(n_0^{(i)},\ \cdots,n_k^{(i)})$ is a nonsingular point, and so (4.27) is a normal point in the Padé-Hermite table for (A(z),B(z)). Thus, according to (4.13), the PHFo $(P^{(i)}(z),Q^{(i)}(z))$ of type (4.27) satisfies

$$A(z)P^{(i)}(z) + B(z)\cdot Q^{(i)}(z) = z^{\mu}R^{(i)}(z), \tag{4.32}$$

where $R^{(i)}(0) = 1$ and

$$\mu = m_0^{(i)} + \cdots + m_k^{(i)} + k$$

$$= n_0^{(i)} + \cdots + n_a^{(i)} + n_d^{(i)} + \cdots + n_k^{(i)} + a-d+k+1. \tag{4.33}$$

Furthermore, the WPHFr $(U^{(i)}(z),V^{(i)}(z))$ of type

$$(n_0^{(i)} + 1, \cdots, n_a^{(i)} + 1, 0, \cdots ,0, n_d^{(i)} + 1, \cdots ,n_k^{(i)}+1) \tag{4.34}$$

satisfies

$$A(z)U^{(i)}(z) + B(z)\cdot V^{(i)}(z) = z^{\mu+2}W^{(i)}(z), \tag{4.35}$$

where $V^{(i)}(0) = I$.

It is required to show that (4.29) is a normal point in the Padé-Hermite table for (A(z),B(z)) if and only if (4.31) is a normal point in the Padé-Hermite table for $(R^{(i)}(z),W^{(i)}(z))$. The proof proceeds by examining the PHFo's and WPHFo's at these two points.

For $s \geq 1$, let $(P'(z),Q'(z))$ be a PHFo of type (4.31) for $(R^{(i)}(z),W^{(i)}(z))$ and let $(U'(z),V'(z))$ be a WPHFo of type

$$(s-1,s, \cdots ,s,n_{a+1}^{(i)}+s+1, \cdots ,n_b^{(i)}+s+1,0, \cdots ,0,n_c^{(i)}+s+1, \cdots ,n_{d-1}^{(i)}+s+1,s, \cdots ,s) \tag{4.36}$$

for $(R^{(i)}(z),W^{(i)}(z))$. Then,

$$R^{(i)}(z)\cdot P'(z) + W^{(i)}(z)Q'(z) = z^{\nu}R'(z) \tag{4.37}$$

$$R^{(i)}(z)\cdot U'(z) + W^{(i)}(z)V'(z) = z^{\nu+2}W'(z), \tag{4.38}$$

where

$$\nu = n_{a+1}^{(i)}+ \cdots +n_b^{(i)}+n_c^{(i)}+ \cdots +n_{d-1}^{(i)}+(b-c+k+2)\cdot s-(a-b+c-d+2). \tag{4.39}$$

Note that $Q'(z) \neq 0$ and $V'(z) \neq 0$, because $R^{(i)}(0) \neq 0$ and either $a \geq 1$ or $d \leq k$. Let

$$P^*(z) = U^{(i)}(z)\cdot Q'(z) + z^2P^{(i)}(z)\cdot P'(z), \tag{4.40}$$

$$Q^*(z) = V^{(i)}(z)\cdot Q'(z) + z^2Q^{(i)}(z)\cdot P'(z), \tag{4.41}$$

$$U^{*}(z) = U^{(i)}(z)\cdot V'(z) + z^{2}P^{(i)}(z)\cdot U'(z) \tag{4.42}$$

and

$$V^{*}(z) = V^{(i)}(z)\cdot V'(z) + z^{2}Q^{(i)}(z)\cdot U'(z). \tag{4.43}$$

We shall show that $(P^{*}(z),Q^{*}(z))$ is a PHFo of type (4.29) for (A(z),B(z)) and that $(U^{*}(z),V^{*}(z))$ is a WPHFo of type

$$(n_0^{(i)}+s+1, \cdots ,n_b^{(i)}+s+1,0, \cdots ,0, n_c^{(i)}+s+1, \cdots ,n_k^{(i)}+s+1) \tag{4.44}$$

for (A(z),B(z)). It is a trivial matter to show that $P^{*}(z)$, $Q^{*}(z)$, $U^{*}(z)$ and $V^{*}(z)$ have correct degree, that is, that condition I in Definitions 3.1 and 3.2 for PHFo's and WPHFo's is satisfied.

Furthermore, because $R^{(i)}(0) \neq 0$, $Q'(z) \neq 0$, $V'(z) \neq 0$ and $V^{(i)}(0) = I$, it follows that $Q^{*}(z) \neq 0$ and $V^{*}(z) \neq 0$, so that condition III in Definitions 3.1 and 3.2 is satisfied. More specifically, suppose that $Q^{*}(z) = 0$ in (4.41). Then

$$Q'(z) = -z^{2}\cdot[V^{(i)}(z)]^{-1}\cdot Q^{(i)}(z)\cdot P'(z). \tag{4.45}$$

Since $Q'(z) \neq 0$, then also $P'(z) \neq 0$. Substitution of (4.45) into (4.37) yields

$$\{R^{(i)}(z) -z^{2}\cdot W^{(i)}(z)\cdot[V^{(i)}(z)]^{-1}\cdot Q^{(i)}(z)\}\cdot P'(z) = z^{\nu}R'(z). \tag{4.46}$$

However, from (4.31) and (4.39), it is seen that

$$\partial(P'(z)) \leq s-2 < \nu. \tag{4.47}$$

Then (4.46) and (4.47) imply that $R^{(i)}(0) = 0$, which is a contradiction. Thus, $Q^{*}(z) \neq 0$ and similarly $V^{*}(z) \neq 0$.

It remains to verify that $(P^{*}(z),Q^{*}(z)$ and $(U^{*}(z),V^{*}(z)$ satisfy condition II of Definitions 3.1 and 3.2, respectively; that is, it remains to verify that

$$A(z)\cdot P^{*}(z) + B(z)\cdot Q^{*}(z) = z^{\omega}\cdot R^{*}(z) \tag{4.48}$$

and

$$A(z)\cdot U^*(z) + B(z)\cdot V^*(z) = z^{\omega+2}\cdot W^*(z), \tag{4.49}$$

where

$$\omega = n_{a+1}^{(i)} + \cdots + n_b^{(i)} + n_c^{(i)} + \cdots + n_{d-1}^{(i)} + (b-c+k+2)\cdot s + (b-c+k+1). \tag{4.50}$$

Using (4.32), (4.35), (4.37), (4.40) and (4.41) we obtain

$$\begin{aligned} A(z)P^*(z) + B(z)Q^*(z) &= A(z)[U^{(i)}(z)Q'(z) + z^2\cdot P^{(i)}(z)P'(z)] + B(z)[V^{(i)}(z)Q'(z) + z^2\cdot Q^{(i)}(z)P'(z)] \\ &= A(z)U^{(i)}(z) + B(z)V^{(i)}(z)Q'(z) + z^2A(z)P^{(i)}(z) + B(z)Q^{(i)}(z)P'(z) \\ &= z^{\nu+2}[R^{(i)}(z)P'(z) + W^{(i)}(z)Q'(z)] \\ &= z^{\mu+\nu+2}R'(z) \\ &= z^{\omega}\cdot R'(z). \end{aligned} \tag{4.51}$$

Thus,

$$R^*(z) = R'(z). \tag{4.52}$$

Similarly, using (4.32), (4.35), (4.38), (4.42) and (4.43), we obtain

$$\begin{aligned} A(z)U^*(z) + B(z)V^*(z) &= A(z)[U^{(i)}(z)V'(z) + z^2\cdot P^{(i)}(z)U'(z)] + B(z)[V^{(i)}(z)V'(z) + z^2\cdot Q^{(i)}(z)U'(z)] \\ &= A(z)U^{(i)}(z) + B(z)V^{(i)}(z)V'(z) + z^2A(z)P^{(i)}(z) + B(z)Q^{(i)}(z)U'(z) \\ &= z^{\nu+2}[R^{(i)}(z)U'(z) + W^{(i)}(z)V'(z)] \\ &= z^{\mu+\nu+4}W'(z) \\ &= z^{\omega+2}\cdot W'(z). \end{aligned} \tag{4.53}$$

To conclude the proof of the theorem, from (4.43) and (4.52), observe that $R^*(0) \neq 0$ and $V^*(0)$ is nonsingular if and only if $R'(0) \neq 0$ and $V'(0)$ is nonsingular. From Theorem 3.6, we can then conclude that (4.29) is a normal point in the Padé-Hermite table for (A(z),B(z)) if and only if (4.31) is a normal point in the Padé-Hermite table for $(R^{(i)}(z),W^{(i)}(z))$.

■

Theorem 4.7: The cofactor Padé-Hermite sequence along with the associated weak Padé-Hermite cofactor sequence for (A(z),B(z)) satisfy

$$\begin{bmatrix} U^{(i+1)}(z) & P^{(i+1)}(z) \\ V^{(i+1)}(z) & Q^{(i+1)}(z) \end{bmatrix} = \begin{bmatrix} U^{(i)}(z) & P^{(i)}(z) \\ V^{(i)}(z) & Q^{(i)}(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix} \tag{4.54}$$

where $(P'(z),Q'(z))$ is the PHFo for $(R^{(i)}(z),W^{(i)}(z))$ of type $(m'_0, \cdots, m'_k)$ which represents the first nonsingular point $(s_i-2,\ s_i-1-\Delta_1^{(i)},\ \cdots,\ s_i-1-\Delta_k^{(i)})$ and $(U'(z),V'(z))$ is its associated WPHFr.

Proof: Since $(P^{(i)}(z),Q^{(i)}(z))$ and $(P^{(i+1)}(z),Q^{(i+1)}(z))$ are successive elements of the cofactor sequence (4.15), then, according to (4.7), (4.9) and (4.10), $(n_0^{(i)}, \cdots, n_k^{(i)})$ and $(n_0^{(i+1)}, \cdots, n_k^{(i+1)})$ are successive nonsingular points along the offdiagonal path starting at $(n_0^{(0)}, \cdots, n_k^{(0)})$. By Theorem 4.6, then s_i is the smallest positive integer for which $(s_i-2, s_i-1-\Delta_1^{(i)},\ \cdots,\ s_i-1-\Delta_k^{(i)})$ is a nonsingular point for $(R^{(i)}(z),W^{(i)}(z))$. Accordingly, we can determine $(P'(z),Q'(z))$ to be a PHFo of type $(m'_0, \cdots, m'_k)$, where $(m'_0, \cdots, m'_k)$ is the representative of the point $(s_i-2, s_i-1-\Delta_1^{(i)},\ \cdots,\ s_i-1-\Delta_k^{(i)})$ in the Padé-Hermite table of $(R^{(i)}(z),W^{(i)}(z))$. The leading term of the residual satisfies $R'(0) = 1$. In addition, we can determine its associated WPHFr $(U'(z),V'(z))$, where $V'(0) = I$.

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

$$\begin{bmatrix} U^*(z) & P^*(z) \\ V^*(z) & Q^*(z) \end{bmatrix} = \begin{bmatrix} U^{(i)}(z) & P^{(i)}(z) \\ V^{(i)}(z) & Q^{(i)}(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix}. \tag{4.55}$$

Using the same arguments as in Theorem 4.6 it is clear that $(P^*(z),Q^*(z))$ is a PHFo of type $(n_0^{(i)}+s_i,\ \cdots,\ n_k^{(i)}+s_i)$. In addition, the leading term of the residual is one. Because of the uniqueness of Padé-Hermite residual and cofactor sequences, $(P^*(z),Q^*(z))$ is the (i+1)-st term of the cofactor sequence. Similarly, $(U^*(z),V^*(z))$ can be shown to be the unique (i+1)-st term of the weak Padé-Hermite cofactor sequence.

∎

Theorem 4.6 and Theorem 4.7 reduces the problem of determining a PHFo (or a corresponding WPHFo) to two smaller problems: determine a PHFo and an associated WPHFr up to a nonsingular point $(n_0^{(i)}, \cdots, n_k^{(i)})$, and then determine a PHFo (or, a WPHFr) of type $(s-2, s-1-\Delta_1^{(i)}, \cdots, s-1-\Delta_k^{(i)})$. The overhead cost of each step of this iteration scheme is the cost of determining the residual power series plus the cost of combining the solutions, i.e., the cost of multiplying equations (4.40), (4.41), (4.42) and (4.43). This overhead cost is generally an order of magnitude less than the cost of simply solving the linear systems (2.9) or (3.9).

In the special case when $k = 1$, a WHPFr is the same as a Padé fraction. In this case equation (4.16) is given by

$$A(z)U^{(i)}(z) + B(z)\cdot V^{(i)}(z) = z^{n_0^{(i)} + n_1^{(i)} + 3}W^{(i)}(z) \tag{4.56}$$

and $(U^{(i)}(z),V^{(i)}(z))$ is a Padé fraction of type $(n_0^{(i)}+1,n_1^{(i)}+1)$ for (A(z),B(z)). If

$$W^{(i)}(z) = z^{\lambda_i}\cdot\overline{W}(z) \tag{4.57}$$

where $\overline{W}(0) = \overline{w}_0 \neq 0$, then it is possible to show that

$$P^{(i+1)}(z) = z^{\lambda_i}\cdot\overline{w}_0^{-1}\cdot U^{(i)}(z),\ \ Q^{(i+1)}(z) = z^{\lambda_i}\cdot\overline{w}_0^{-1}\cdot V^{(i)}(z),\ \text{ and }\ R^{(i+1)}(z) = \overline{w}_0^{-1}\cdot\overline{W}^{(i)}(z). \tag{4.58}$$

Travelling from one nonsingular point to the next can then be shown to to the same as power series division of one residual into the next.

When $k = 1$, the Extended Euclidean Algorithm for computing polynomial GCD's is closely related to the calculation of Padé approximants (c.f., McEliece and Shearer[11] or Cabay and Choi[3]). When the input power series A(z) and B(z) are polynomials of degree m and n, respectively, then reversing the order of the coefficients in equation (4.56) gives

$$A^*(z)P^{*(i)}(z) + B^*(z)Q^{*(i)}(z) = R^{*(i)}(z). \tag{4.59}$$

Here $A^*(z) = A(z^{-1})z^m, \ \cdots$ etc. Equation (4.59) is similar to the type of equation found in the EEA applied to $(A^*(z),B^*(z))$. In fact, when we are calculating the Padé approximant of type (n,m) for (A(z),B(z)), the reversed residual $R^{*(i)}(z)$ is the i-th term of the remainder sequence calculated in the EEA, while $(P^{*(i)}(z),Q^{*(i)}(z))$ is the i-th term of the cofactor sequence calculated in the EEA. Indeed, this is the primary reason for the naming convention of Definition 4.1 and Definition 4.2.

5. The Algorithm:

Given non-negative integers $(n_0, \cdots, n_k)$, the algorithm PADE_HERMITE below makes use of Theorem 4.7 to compute both the Padé-Hermite and weak Padé-Hermite cofactor sequences (4.15) and (4.18), respectively. Thus, intermediate results available from PADE_HERMITE include those PHFo's for (A(z),B(z)) at all nonsingular points $(n_0^{(i)}, \cdots, n_k^{(i)})$ $i=1,2, \cdots, l-1$, smaller than $(n_0, \cdots, n_k)$ along the straight line path from $(n_0^{(0)}, \cdots, n_k^{(0)})$ to $(n_0, \cdots, n_k)$),together with those WPHFr at the succeeding points. The output gives results associated with the final point $(n_0^{(l)}, \cdots, n_k^{(l)})$. If this final point is a nonsingular point, then the output $(P^{(l)}(z),Q^{(l)}(z))$ is a PHFo having residual beginning with a one, and $(U^{(l)}(z),V^{(l)}(z))$ is the WPHFr at the successor point. If $(n_0, \cdots, n_k)$ is a singular point, then the output $(P^{(l)}(z),Q^{(l)}(z))$ is a PHFo of type $(n_0, \cdots, n_k)$, while $(U^{(l)}(z),V^{(l)}(z))$ is a WPHFo at the successor point. Note that when a WPHFo is required at $(n_0, \cdots, n_k)$, the input to the algorithm should be $(n_0-1, \cdots, n_k-1)$, rather than $(n_0, \cdots, n_k)$.

The algorithm is presented in two parts. The first, INITIAL_PH, takes as its input (1) a power series, R(z), with a nonzero leading term, (2) a $1\times k$ vector of power series $W(z)$, (3) a $(1\times(k+1)$ vector of integers $(n'_0, \cdots, n'_k)$, representing the first location on the offdiagonal line to search for the nonsingular point, and (4) a nonzero integer M representing the maximum distance to travel on the offdiagonal line. The algorithm returns (1) the distance s to the first nonsingular point along the specified offdiagonal line, with $s = M$ if all nodes are singular, (2) the PHFo at this point, and (3) the WPHFo at the successor point.

The main algorithm, PADE_HERMITE calls INITIAL_PH to iteratively construct PHFo's and WPHFo's for the residuals, $(R^{(i)}(z),W^{(i)}(z))$. The PHFo's $(P^{(i)}(z),Q^{(i)}(z))$ and the WPHFo's $(U^{(i)}(z),V^{(i)}(z))$ for (A(z),B(z)) are computed using the results of Theorem 4.7. The validity of the algorithm, therefore, rests with the correctness of Theorem 4.7.

INITIAL_PH(R(z),W(z),$(n'_0, \cdots ,n'_k),M$)

I-1) $s \leftarrow 0$

I-2) $d \leftarrow 0$

I-3) Do while $s < M$ and d $= 0$

I-4) $\quad s \leftarrow s + 1$

I-5) $\quad (m'_0, \cdots , m'_k) \leftarrow (n'_0+s, \cdots , n'_k+s) + (\Delta'_0, \cdots ,\Delta'_k)$,

where $\Delta'_j = -1-n'_j-s$ if $n'_j + s < -1$, and $\Delta'_j = 0$ otherwise

I-6) $\quad$ Compute $d \leftarrow \det(T_{(m'_0, \cdots , m'_k)})$

End do

I-7) Solve (c.f. (2.9))

$$S_{(m'_0, \cdots , m'_k)} \cdot \begin{bmatrix} p_{0,0} \\ \cdot \\ p_{0,m'_0} \\ \cdot \\ p_{k,m'_k} \end{bmatrix} = 0.$$

(if $m'_i < 0$, then $P_i(z) = 0$ and $p_{i,j}$ does not appear on the left side)

I-8) Solve (c.f. (3.9))

$$S'_{(m'_0+1, \cdots , m'_k+1)} \cdot \begin{bmatrix} u_0 \\ \cdot \\ u_{m'_0+1} \\ v_{1,0} \\ \cdot \\ v_{1,m'_1+1} \\ \cdot \\ v_{k,m'_k+1} \end{bmatrix} = 0.$$

I-9) Set

$$P'(z) \leftarrow \sum_{j=0}^{m'_0} p_{0,j} z^j \,, \qquad Q'(z) \leftarrow \begin{bmatrix} \sum_{j=0}^{m'_0} p_{1,j} z^j \\ \cdot \\ \cdot \\ \sum_{j=0}^{m'_0} p_{k,j} z^j \end{bmatrix}$$

I-10) Set

$$U'(z) \leftarrow \sum_{j=0}^{m'_0+1} u_j z^j \,, \qquad V'(z) \leftarrow \begin{bmatrix} \sum_{j=0}^{m'_0+1} v_{1,j} z^j \\ \cdot \\ \cdot \\ \sum_{j=0}^{m'_0+1} v_{k,j} z^j \end{bmatrix}$$

If $d \neq 0$ then normalize :

I-11) Compute $r_0 = R'(0)$ satisfying $R(z) \cdot P'(z) + W(z) \cdot Q'(z) = z^{m'_0 + \cdots + m'_k + k - 1} \cdot R'(z)$

I-12) $P'(z) \leftarrow P'(z) \cdot r_0^{-1}$, $Q'(z) \leftarrow Q'(z) \cdot r_0^{-1}$,

I-13) $U'(z) \leftarrow U'(z) \cdot V'(0)^{-1}$, $V'(z) \leftarrow V'(z) \cdot V'(0)^{-1}$,

End If

I-14) Return$\left(s, \begin{bmatrix} U'(z) & P'(z) \\ V'(z) & Q'(z) \end{bmatrix}\right)$

■

The main algorithm, PADE_HERMITE takes as its input a vector of power series and a vector of integers, each having k+1 components. The vector of integers must have non-negative entries (otherwise one calls PADE_HERMITE with a smaller value of k).

PADE_HERMITE$((A_0(z), \cdots, A_k(z)), (n_0, \cdots, n_k))$

PH-1) Find the largest β such that $A_i(z) = z^{\beta} \cdot \overline{A}_i(z)$ are still power series. Set $A_i(z) = z^{-\beta} \cdot A_i(z)$.

Reorder the power series according to (4.3) and (4.4).

PH-2) $M \leftarrow \begin{cases} n_0 + 1, & \text{if } n_k \geq n_0, \\ max(n_1, n_k) + 2, & otherwise \end{cases}$

PH-3) $(n_0^{(0)}, \cdots, n_k^{(0)}) \leftarrow (n_0 - M, \cdots, n_k - M)$

PH-4) $(s_0, \begin{bmatrix} U^{(1)}(z) & P^{(1)}(z) \\ V^{(1)}(z) & Q^{(1)}(z) \end{bmatrix}) \leftarrow INITIAL\_PH(A(z), B(z), (n_0^{(0)}, \cdots, n_k^{(0)}), M)$

PH-5) $(n_0^{(1)}, \cdots, n_k^{(1)}) \leftarrow (n_0^{(0)} + s_0, \cdots, n_k^{(0)} + s_0)$

PH-6) $i \leftarrow 1$

PH-7) $M \leftarrow M - s_0$

PH-8) Do while $M > 0$

PH-9) $(m_0^{(i)}, \cdots, m_k^{(i)}) \leftarrow (n_0^{(i)}, \cdots, n_k^{(i)}) + (\Delta_0^{(i)}, \cdots, \Delta_k^{(i)})$,

where $\Delta_j^{(i)} = -1 + n_j^{(i)}$ if $n_j^{(i)} < -1$ and $\Delta_j^{(i)} = 0$ otherwise.

PH-10) Determine $R^{(i)}(z)$ from the equation

$$A(z) \cdot P^{(i)}(z) + B(z) \cdot Q^{(i)}(z) = z^{m_0^{(i)} + \cdots + m_k^{(i)} + k} \cdot R^{(i)}(z)$$

PH-11) Determine $W^{(i)}(z)$ from the equation

$$A(z) \cdot U^{(i)}(z) + B(z) \cdot V^{(i)}(z) = z^{m_0^{(i)} + \cdots + m_k^{(i)} + k + 2} \cdot W^{(i)}(z)$$

PH-12) $(s_i, \begin{bmatrix} U'(z) & P'(z) \\ V'(z) & Q'(z) \end{bmatrix}) \leftarrow INITIAL\_PH(R^{(i)}(z), W^{(i)}(z), (-2, -1-\Delta_1^{(i)}, \cdots, -1-\Delta_k^{(i)}), M)$

PH-13)
$$\begin{bmatrix} U^{(i+1)}(z) & P^{(i+1)}(z) \\ V^{(i+1)}(z) & Q^{(i+1)}(z) \end{bmatrix} \leftarrow \begin{bmatrix} U^{(i)}(z) & P^{(i)}(z) \\ V^{(i)}(z) & Q^{(i)}(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix}$$

(c.f. (4.54))

PH-14) $M \leftarrow M - s_i$

PH-15) $(n_0^{(i+1)}, \cdots, n_k^{(i+1)}) \leftarrow (n_0^{(i)} + s_i, \cdots, n_k^{(i)} + s_i)$

PH-16) $i \leftarrow i + 1$

End while

PH-17) Return($\begin{bmatrix} U^{(i+1)}(z) & P^{(i+1)}(z) \\ V^{(i+1)}(z) & Q^{(i+1)}(z) \end{bmatrix}$)

6. Complexity of **Padé_Hermite Algorithm:**

In assessing the cost of PADE_HERMITE, we count the number of multiplications required by most of the steps of the algorithm, excluding from consideration the more trivial ones.

Consider first the cost of invoking the initialization algorithm, INITIAL_PH. Gaussian elimination can be used in step I-6 to obtain a triangular factorization of $T_{(m'_0, \cdots, m'_k)}$. Assuming that the elimination is performed by applying bordering techniques (as s increases), step I-6 requires approximately $(m'_0 + \cdots + m'_k + k + 1)^3/3$ multiplications in F, where $(m'_0, \cdots, m'_k)$ are the values attained upon exit from the WHILE loop I-3. In steps I-7 and I-8, the solutions $(P'(z), Q'(z))$ and $(U'(z), V'(z))$ can then be obtained by forward and backward substitution requiring approximately $(k+1)\cdot(m'_0 + \cdots + m'_k + k + 1)^2$ multiplications in total. Since all invocations of Initial_PH specify $n_i \leq -1$, the WHILE loop I-3 yields $m'_i \leq s$, for $0 \leq i \leq k$, where s is the step size. Consequently, for this step size s, the total cost of INITIAL_PH is

$$Cost(INITIAL\_PH) \approx (k+1)^3(s+1)^3 + (k+1)^3\cdot(s+1)^2$$

$$\approx (k+1)^3(s+1)^3. \quad (6.1)$$

For the main routine PADE_HERMITE, the approximate costs (in terms of multiplications) associated with the major steps are summarized in Table 6.1. In the table, we assign

$$\lambda_i = m_0^{(i)} + \cdots + m_k^{(i)} + k + 1. \quad (6.2)$$

Variable Computed	Step	Approximate Cost
$P^{(1)}, Q^{(1)}, U^{(1)}, V^{(1)}$	PH-4	$(k+1)^3\cdot(s_0+1)^3$
$R^{(i)}$	PH-10	$\lambda_i\cdot(k+1)\cdot(s_i+1)$
$W^{(i)}$	PH-11	$(\lambda_i+k+1)\cdot(k+1)^2\cdot(s_i+1)$
P', Q', U', V'	PH-12	$(k+1)^3\cdot(s_i+1)^3$
$P^{(i+1)}, Q^{(i+1)}$	PH-13	$\lambda_i\cdot(k+1)(s_i-1)$
$U^{(i+1)}, V^{(i+1)}$	PH-14	$\lambda_i\cdot(k+1)^2 s_i$

Table 6.1

Bounds on Multiplications per step

In steps PH-10 and PH-11, it is assumed that $R^{(i)}(z)$ and $W^{(i)}(z)$ are computed only to the number of terms required to obtain the next s_i. This can be accomplished, for example, by declaring the power series passed to INITIAL_PH and the integer s returned by it to be global variables.

Using Table 6.1, we obtain

Theorem 6.1. The algorithm PADE_HERMITE requires

$$O((k+1)^2(n_0^2 + \cdots + n_k^2)) + O((k+1)^3\varsigma^2\eta) \tag{6.3}$$

multiplications in F, where

$$\varsigma = max(s_0, s_1, \cdots) \ , \ \text{and} \quad \eta = max(n_0, \cdots , n_k). \tag{6.4}$$

In particular, the algorithm requires

$$O((k+1)^2(n_0^2 + \cdots + n_k^2)) \tag{6.5}$$

multiplications in the normal case.

Proof: Let l be the number of steps required in a PADE_HERMITE calculation. Then, it is easy to show that

$$\sum_{i=0}^{l} s_i = \eta \tag{6.6}$$

and

$$\sum_{i=0}^{l} s_i(m_j^{(i)} + 1) \le n_j^2/2. \tag{6.7}$$

Thus,

$$\begin{aligned} \sum \lambda_i s_i &= \sum_{i=0}^{l} s_i \cdot (m_0^{(i)} + \cdots + m_k^{(i)} + k + 1) \\ &= \sum_{j=0}^{k}\sum_{i=0}^{l} s_i(m_j^{(i)} + 1) \\ &\le \sum_{j=0}^{k} n_j^2/2. \end{aligned} \tag{6.8}$$

Also

$$\sum_{i=0}^{l} (k+1)^3 s_i^3 \le (k+1)^3 \varsigma^2 \sum_{i=0}^{l} s_i \le (k+1)^3 \cdot \varsigma^2 \eta. \tag{6.9}$$

Equations (6.8) and (6.9) along with Table 6.1 complete the proof.

■

When $n_0 = \cdots = n_k = n$, the complexity of PADE_HERMITE in the normal case is $O((k+1)^3 \cdot n^2)$. If $N = (k+1) \cdot n$ is the size of the associated Sylvester matrix, then this says that the system (2.9) can be solved using $O((k+1) \cdot N^2)$ operations. This agrees with the results of Kailath et al [8] under the same normality assumptions. In the nonnormal case, however, their algorithm breaks down and so a method such as Gaussian elimination, with a cost of $O((k+1)^3 \cdot n^3)$ operations, is required. With the use of PADE_HERMITE however, even the existence of only one nonsingular point along the offdiagonal results in significant speedup. For example, if the point $(n/2, \cdots, n/2)$ is the only nonsingular point on the main offdiagonal, then the cost of determining a PHFo of type $(n, \cdots, n)$ is reduced by a factor of 4.

Under the assumption of normality, the algorithm PADE_HERMITE has approximately the same complexity as present efficient algorithms. However, when it is known a priori that the normality assumption can be made, then we are no longer constrained in our step size. In particular, the algorithm may be altered to have a step size of 2^i at the i-th step. If M is the distance to be traveled along the given off-diagonal in $k+1$ space, then the altered algorithm will require at most log n steps. By stepping forward in powers of two, the residuals can be determined using fast multiplication techniques. In the special case where $n_0 = \cdots = n_k = n$, the result will be an algorithm that constructs a Padé-Hermite form with a cost of $O((k+1) \cdot n \cdot \log^2 n)$ operations. In a similar manner, the assumption of normality used in conjunction with the recurrence relation (4.54), allows for the use of a recursive algorithm based on a divide-and-conquer approach. Again, under the assumption of normality, one obtains an algorithm that is faster than existing methods. That this speedup can also be accomplished without the assumption of normality is an open question.

References

1. G.A. Baker, *Essentials of Padé Approximants,* Academic Press, New York (1975).

2. M.G. De Bruin, "Some convergence results in simultaneous rational approximation to the set of hypergeometric functions $\{F_1(1;c_i,z)\}$," pp. 95-117 in *Padé Seminar 1983*, ed. H.J. Bunger, (1983).

3. S. Cabay and D.K. Choi, "Algebraic Computations of Scaled Padé Fractions," *SIAM J. of Computing,* pp. 243-270 (Feb, 1986).

4. J. Della Dora and C. Dicrescenzo, "Approximants de Padé-Hermite," *Numer. Math,* **43** pp. 23-57 (1984).

5. W.B. Gragg, "The Padé Table and its Relation to Certain Algorithms of Numerical Analysis," *SIAM Rev.*, **14** pp. 1-61 (1972).

6. P. Graves-Morris, "Vector Rational Interpolants II," *IMA Journal of Numerical Analysis*, **4** pp. 209-224 (1984).

7. M.Ch. Hermite, "Extrait d'une lettre de M. Ch. Hermite a Paul Gordan," *J. de Crelle*, **76** pp. 303-311 (1873).

8. T. Kailath, S-Y. Kung, and M. Morf, "Displacement Ranks of Matrices and Linear Equations," *J. Math. Anal. Appl.*, **68** pp. 395-407 (1979).

9. G. Labahn and S. Cabay, "Matrix Padé Fractions and their Computation," *SIAM J. of Computing*, p. 19 pgs (June, 1989).

10. G. Labahn, D.K. Choi, and S. Cabay, "Inverses of Block Hankel and Block Toeplitz Matrices," *SIAM J. of Computing*, (to appear).

11. R.J. McEliece and J.B. Shearer, "A Property of Euclid's Algorithm and an Application to Padé Approximation," *SIAM J. Appl. Math*, **34** pp. 611-617 (1978).

12. H. Padé, "Sur la Representation Approchee d'une Fonction par des Fractions Rationelles," *Ann Ecole Nor.*, pp. 1-93, supplement Thesis, (1892).

13. S. Paszkowski, "Recurrence Relations in Padé-Hermite Approximation," *J. of Comp. and Appl. Math*, **19** pp. 99-107 (1987).

14. R.E. Shafer, "On Quadratic Approximation," *SIAM J. Numer. Anal.*, **11** pp. 447-460 (1974).