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

We present a new fast algorithm for the calculation of a Padé-Hermite form for a vector of power series. 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.

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 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 [11] in his classic thesis, we wish to select k+1 polynomials so that for y(z) as the given power series we have

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

where N is the sum of the degrees, $\{n_i\}$, 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

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

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

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

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 [13] 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 [12].

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 [12]). 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 traveling 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. In the normal case, however, the iteration scheme can be modified to give an alternate algorithm. The resulting scheme is then more efficient than existing methods. A brief sketch of such an algorithm is also included.

2. Basic Definitions

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

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

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

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) $$\sum_{i=0}^{k} A_i(z) P_i(z) = z^{n_0 + \cdots + n_k + k} \cdot R(z), \tag{2.2}$$
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 $A_1(z) = -1$, 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 [13].

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

$$A^2(z)P(z) + Q(z) = O(z^{m+n+1}), \tag{2.3}$$

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).

When $A_i(0) = 0$, $0 \leq i \leq k$, we can remove the largest factor, z^β from all the power series. Any PHFo of type $(n_0, \cdots, n_k)$ for $(z^{-\beta}A_0(z), \cdots, z^{-\beta}A_k(z))$ is then also a PHFo of the same type for $(A_0(z), \cdots, A_k(z))$. Thus, for algorithmic purposes we may assume that $A_l(0) \neq 0$ for at least one l. By renumbering if necessary, we will henceforth assume that $a_{0,0} = A_0(0) \neq 0$. In this case, we will also set

$$A(z) = A_0(z),\; B(z) = (A_1(z), \cdots, A_k(z)). \tag{2.4}$$

We view B(z) as a $1 \times k$ matrix of power series. Note that, using (2.4), equation (2.2) can be written in a matrix format as

$$A(z) \cdot P(z) + B(z) \cdot Q(z) = z^{n_0 + \cdots + n_k + k} \cdot R(z), \tag{2.5}$$

where $P(z) = P_0(z)$, and $Q(z) = [P_1(z), \cdots, P_k(z)]^t$.

Central to our approach for efficiently calculating PHFo's is a second type of generalization of a Padé form.

Definition 2.2. 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) $$A(z) \cdot U(z) + B(z) \cdot V(z) = z^{n_0 + \cdots + n_k + 1} \cdot W(z), \tag{2.6}$$
where W(z) is a $1 \times k$ matrix of power series, and

III) the columns of V(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 the constant term, V(0), of the weak Padé-Hermite denominator is a nonsingular matrix, then we say that (U(z),V(z)) is a **weak Padé-Hermite fraction (WPHFr).**

When $k = 1$, Definition 2.2 is the scalar definition of a Padé form (c.f., Gragg [5]) while a WPHFr is equivalent to a scalar Padé fraction. A WPHFr can be interpreted as providing a a set of simultaneous Padé approximants for the quotient power series $A_i(z)/A_0(z)$ (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

$$B(z)/A_0(z) \approx -U(z)\cdot V(z)^{-1}. \tag{2.7}$$

Since

$$U(z)\cdot V(z)^{-1} = U(z)\cdot adj(V(z))/det(V(z)), \tag{2.8}$$

equations (2.7) and (2.8) give a simultaneous rational approximation for each power series

$$\frac{A_i(z)}{A_0(z)} \approx \frac{N_i(z)}{D(z)},\ i = 1, .., k. \tag{2.9}$$

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]).

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

$$P(z) = p_0 + p_1 z + \cdots + p_n z^n, \tag{2.10}$$

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

$$P = [p_0, \cdots, p_n]^t. \tag{2.11}$$

Let $S_{(n_0, \cdots, n_k)}$ be the matrix

$$\begin{bmatrix} a_{0,0} & & & | & & | & a_{k,0} & & \\ & & & | & & | & & \cdot & \\ \cdot & & & & & & & & a_{k,0} \\ \cdot & & a_{0,0} & | & \cdots & | & \cdot & & \cdot \\ \cdot & & \cdot & & & & \cdot & & \\ & & \cdot & | & & | & & & \\ a_{0,\lambda} & \cdot & a_{0,\lambda-n_0} & & & & a_{k,\lambda} & \cdot & a_{k,\lambda-n_k} \end{bmatrix}, \tag{2.12}$$

where

$$\lambda = n_0 + \cdots + n_k + k - 1. \tag{2.13}$$

$S_{(n_0, \cdots, n_k)}$ denotes 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 (2.2) can be written as

$$S_{(n_0, \cdots, n_k)} \begin{bmatrix} P_0 \\ \cdot \\ \cdot \\ P_k \end{bmatrix} = 0; \tag{2.14}$$

whereas equation (2.6) can be written as

$$S^*_{(n_0, \cdots, n_k)} \begin{bmatrix} U \\ V_1 \\ \cdot \\ \cdot \\ V_k \end{bmatrix} = 0. \tag{2.15}$$

In (2.15), $S^*_{(n_0, \cdots, n_k)}$ is the matrix obtained by deleting the last $(k-1)$ rows from $S_{(n_0, \cdots, n_k)}$ and $V_i(z)$ represents the i-th row of V(z). Equation (2.15) accounts for the naming convention used, since it shows that a WPHFo is the same as a PHFo, except for a weaking of the linear system that defines it. Equations (2.14) and (2.15) imply

Theorem 2.3 (Existence of PHFo's and WPHFo's): Let (A(z),B(z)) be as in (2.4). 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

1) a PHFo of type $(n_0, \cdots, n_k)$ for (A(z),B(z)).
2) a WPHFo of type $(n_0+1, \cdots, n_k+1)$ for (A(z),B(z)). ∎

From Theorem 2.3, it follows that there are many possible choices for a PHFo or a WPHFo of a given type. As in the case of Padé approximants, it is desirable to have PHFo's and WPHFo's that are unique, at least up to a multiplicative constant.

With λ defined by (2.13), let $T_{(n_0, \cdots, n_k)}$ be the $(\lambda+2) \times (\lambda+2)$ matrix whose first $\lambda+1$ rows come from $S_{(n_0, \cdots, n_k)}$ and with last row given by

$$[a_{0,\lambda+1}, \cdots, a_{0,\lambda-n_0+1}, \cdots, a_{k,\lambda+1}, \cdots, a_{k,\lambda-n_k+1}]. \tag{2.16}$$

In addition, set

$$d_{(n_0, \cdots, n_k)} = det(T_{(n_0, \cdots, n_k)}). \tag{2.17}$$

When $n_i = 0$ for all i, we define $d_{(n_0, \cdots, n_k)} = 0$.

It is not hard to show that the nonsingularity of the generalized Sylvester matrix (2.16) results in PHFo's that are unique up to multiplication by a nonzero scalar and in WPHFo's that are unique up to multiplication on the right by nonsingular $k \times k$ matrices. What is surprising, however, is that the existence and uniqueness of a PHFo and a WPHFo satisfying certain extra constraints give both necessary and sufficient conditions for the nonsingularity of these matrices. Indeed, we have

Theorem 2.4. Let (A(z),B(z)) be as in (2.4) and let $(n_0, \cdots, n_k)$, $n_i \ge -1$, be a vector of integers. Then $d_{(n_0, \cdots, n_k)} \ne 0$ if and only if

1) a PHFo (P(z),Q(z)) of type $(n_0, \cdots, n_k)$ is unique up to multiplication by a nonzero constant. In addition, the leading term, R(0), of the residual in condition II of PHFo is nonzero.
2) there exists a **WPHFr** of type $(n_0+1, \cdots, n_k+1)$ for

(A(z),B(z)) unique up to multiplication on the right by a nonsingular $k \times k$ matrix.

∎

Theorem 2.4 is central to the results in this paper.

3. Padé-Hermite Residual Sequences

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[12]) 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). In the non-normal case, individual points $(n_0, \cdots, n_k)$ having the property that $d_{(n_0, \cdots, n_k)} \neq 0$ will be called **nonsingular** points of the vector of power series.

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 (2.14) 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 (2.15). However, neither of these calculations 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 the systems (2.14) and (2.15), 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

$$A_0(0) \neq 0, \ \cdots, A_l(0) \neq 0, \ A_j(0) = 0, \ \text{for } j > l \tag{3.3}$$

and

$$n_0 \geq \cdots \geq n_l, \ \text{and} \ n_{l+1} \leq \cdots \leq n_k. \tag{3.4}$$

Also, let

$$M = \begin{cases} n_0 + 1, & \text{if } n_k \geq n_0, \\ max(n_1, n_k) + 2, & otherwise. \end{cases} \tag{3.5}$$

and

$$(n_0^{(0)}, \cdots, n_k^{(0)}) = (n_0 - M, n_1 - M, \cdots, n_k - M) \tag{3.6}$$

Along the line in (k+1)-space from $(n_0^{(0)}, \cdots, n_k^{(0)})$ to $(n_0, \cdots, n_k)$ define a sequence of points

$$(n_0^{(1)}, \cdots, n_k^{(1)}), \ (n_0^{(2)}, \cdots, n_k^{(2)}), \ \cdots \tag{3.7}$$

by letting $(n_0^{(i)}, \cdots, n_k^{(i)})$ be the i-th nonsingular point along this line.

Corresponding to the sequence (3.7), we introduce

$$(m_0^{(1)}, \cdots, m_k^{(1)}), \ (m_0^{(2)}, \cdots, m_k^{(2)}), \ \cdots \tag{3.8}$$

where

$$m_j^{(i)} = n_j^{(i)} \ \ \text{if } n_j^{(i)} \geq -1 \tag{3.9}$$

and -1 otherwise.

For $i = 1, 2, \ldots$, 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 2.4, there exists a power series $R^{(i)}(z)$ such that

$$A(z)P^{(i)}(z) + B(z) \cdot Q^{(i)}(z) = z^{M^{(i)} + k} R^{(i)}(z) \tag{3.10}$$

where $M^{(i)} = m_0^{(i)} + \cdots + m_k^{(i)}$ and $R^{(i)}(0) \neq 0$. This PHFo is made unique by insisting that $R^{(i)}(0) = 1$ (c.f. Theorem 2.4).

Definition 3.1 The sequence

$$\{R^{(i)}(z)\}, \ i = 1, 2, \cdots, \tag{3.11}$$

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

$$\{(P^{(i)}(z), Q^{(i)}(z))\}, \ i = 1, 2, \cdots, \tag{3.12}$$

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 $1 \times k$ matrix of power series $W^{(i)}(z)$ such that

$$A(z)U^{(i)}(z) + B(z) \cdot V^{(i)}(z) = z^{M^{(i)} + k + 2} W^{(i)}(z) \tag{3.13}$$

where $det(V^{(i)}(0)) \neq 0$. This WPHFr is made unique by insisting that $V^{(i)}(0) = I$ (c.f. Theorem 2.4).

Definition 3.2. The sequence

$$\{W^{(i)}(z)\}, \ i = 1, 2, \cdots, \tag{3.14}$$

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

$$\{(U^{(i)}(z), V^{(i)}(z))\}, \ i = 1, 2, \cdots, \tag{3.15}$$

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

∎

The algorithm described in Section 4 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 3.3 below gives 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

$$\Delta_j^{(i)} = -1 - n_j^{(i)} \ \ \text{if } n_j^{(i)} < -1 \tag{3.16}$$

and 0 otherwise. The following theorem represents main result of this section. Its proof depends heavily on the easy recognition of nonsingular points given in Theorem 2.4.

Theorem 3.3: The cofactor Padé-Hermite sequence along with the associated weak Padé-Hermite cofactor sequence for (A(z),B(z)) satisfy

$$\begin{bmatrix} U^{(i+1)}(z) & P^{(i+1)}(z) \\ V^{(i+1)}(z) & Q^{(i+1)}(z) \end{bmatrix} = \begin{bmatrix} U^{(i)}(z) & P^{(i)}(z) \\ V^{(i)}(z) & Q^{(i)}(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix} \quad (3.17)$$

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 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.

∎

Theorem 3.3 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 $(s_i-2, s_i-1-\Delta_1^{(i)}, \cdots, s_i-1-\Delta_k^{(i)})$. The overhead cost of each step of this iteration scheme is the cost of determining the residual power series, along with the cost of combining the solutions, (i.e., the cost of multiplying out equation (3.17)). This overhead cost is generally an order of magnitude less than the cost of simply solving the linear systems (2.14) or (2.15).

In the special case when $k = 1$, a WPHFr is the same as a Padé fraction. In this case equation (3.13) is the same as

$$A_0(z)\cdot U^{(i)}(z) + A_1(z)\cdot V^{(i)}(z) = z^{m^{(i)}+n^{(i)}+3}\cdot W^{(i)}(z) \quad (3.18)$$

and $(U^{(i)}(z),V^{(i)}(z))$ is a Padé fraction of type $(m^{(i)}+1,n^{(i)}+1)$. If

$$W^{(i)}(z) = z^{\lambda_i}\hat{W}^{(i)}(z) \quad (3.19)$$

with $\hat{W}(0) = \hat{w}_0 \neq 0$, then by the uniqueness of the cofactor sequences it is easy to show that

$$P^{(i+1)}(z) = z^{\lambda_i}\hat{w}_0^{-1}U^{(i)}(z),\ Q^{(i+1)}(z) = z^{\lambda_i}\hat{w}_0^{-1}V^{(i)}(z), \quad (3.20)$$

and

$$R^{(i+1)}(z) = \hat{w}_0^{-1}\cdot\hat{W}(z). \quad (3.21)$$

Traveling from one nonsingular point to the next can then be shown to be 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 Padé approximation (c.f., McEliece and Shearer [10] or Cabay and Choi [3]). When, in addition, the input power series $A_0(z)$ and $A_1(z)$ are polynomials of degree m and n, respectively, then reversing the coefficients in equation (3.18) gives

$$A^*_0(z)\cdot P^{*(i)}(z) + A^*_1(z)\cdot Q^{*(i)}(z) = R^{*(i)}(z). \quad (3.22)$$

Here $A^*_0(z) = A_0(z^{-1})\cdot z^m, \cdots$ etc.. Equation (3.22) is similar to the type of equation found in the EEA applied to $(A^*_0(z),A^*_1(z))$. In fact, when we are calculating the Padé approximant of type (n,m), $R^{*(i)}(z)$ 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 corresponding cofactor sequence calculated in the EEA. Indeed, this is the primary reason for the naming convention of Definitions 3.1 and 3.2.

4. The Algorithm:

From the recurrence relation given in Theorem 3.3 we construct an algorithm, PADE_HERMITE, which can be summarized as follows.

Algorithm PADE_HERMITE:

Input: A vector of power series $(A_0(z), \cdots, A_k(z)) = (A(z),B(z))$ and a vector of nonnegative integers $(n_0, \cdots, n_k)$.
Output: A Padé-Hermite form of type $(n_0, \cdots, n_k)$ for (A(z),B(z)).

Step#1: Find the first nonsingular point along the $(n_0, \cdots, n_k)$ off-diagonal in (k+1)-space. If this point is at $(n_0^{(1)}, \cdots, n_k^{(1)})$, then return the unique PHFo of this type, along with the corresponding unique WPHFr associated with this form. This is done in a subroutine, INITIAL_PADE_HERMITE.

Step#2: Determine the residuals of both the PHFo and the WPHFr at the present nonsingular point.

Step#3: Using INITIAL_PADE_HERMITE, determine the first nonsingular point for the residuals along the line specified by Theorem 3.3.

Step#4: Combine the results (using Theorem 3.3) to obtain a PHFo and WPHFr at the next nonsingular point.

Step#5: Either terminate if successful, or iterate back to step #2.

∎

The complexity of PADE_HERMITE can be summarized by

Theorem 4.1. The algorithm PADE_HERMITE requires

$$O((k+1)^2(n_0^2 + \cdots + n_k^2)) + O((k+1)^3\varsigma^2\eta) \quad (4.1)$$

multiplications in F, where

$$\varsigma = max(s_0, s_1, \cdots),\ \text{and}\ \eta = max(n_0, \cdots, n_k). \quad (4.2)$$

In particular, the algorithm requires

$$O((k+1)^2(n_0^2 + \cdots + n_k^2)) \quad (4.3)$$

multiplications in the normal case.

∎

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 (2.14) 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 an approximate factor of 4 (since we are solving two systems, each having a cost of $O((k+1)^2 \cdot (n/2)^3)$ operations).

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. 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).
11. H. Padé, "Sur la Representation Approchee d'une Fonction par des Fractions Rationelles," *Ann Ecole Nor.,* pp. 1-93, supplement Thesis, (1892).
12. S. Paszkowski, "Recurrence Relations in Padé-Hermite Approximation," *J. of Comp. and Appl. Math,* **19** pp. 99-107 (1987).
13. R.E. Shafer, "On Quadratic Approximation," *SIAM J. Numer. Anal.,* **11** pp. 447-460 (1974).