Journal of Computational and Applied Mathematics 39 (1992) 295–313
North-Holland

On the theory and computation of nonperfect Padé–Hermite approximants

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

Bernhard Beckermann
Institut für Angewandte Mathematik, Universität Hannover, Welfengarten 1, W-3000 Hannover, Germany

Received 24 August 1990
Revised 8 December 1990

Abstract

Cabay, S., G. Labahn and B. Beckermann, On the theory and computation of nonperfect Padé–Hermite approximants, Journal of Computational and Applied Mathematics 39 (1992) 295–313.

For a vector of $k+1$ power series we introduce two new types of rational approximations, the weak Padé–Hermite form and the weak Padé–Hermite fraction. A recurrence relation is then presented which computes Padé–Hermite forms along with their weak counterparts along a sequence of perfect points in the Padé–Hermite table. The recurrence relation results in a fast algorithm for calculating a Padé–Hermite approximant of any given type. When the vector of power series is perfect, the algorithm is shown to calculate a Padé–Hermite form of type $(n_0, \ldots, n_k)$ in $O(kN)^2$ operations, where $N = n_0 + \cdots + n_k$. This complexity is the same as that of other fast algorithms. The new algorithm also succeeds in the nonperfect case, usually with only a moderate increase in cost.

Keywords: Vector of power series, Padé–Hermite fraction, Padé–Hermite approximation, rational approximation, Sylvester matrix.

1. Introduction

Given a vector of $k+1$ power series

$$A_i(z) = \sum_{j=0}^{\infty} a_{i,j} z^j, \tag{1.1}$$

* Supported in part by NSERC grant No. A8035.
** Supported in part by NSERC grant No. FS1525C.

with coefficients from a field, a Padé–Hermite approximant of type $(n_0,\ldots, n_k)$ is a set of $k+1$ polynomials $P_i(z)$ having degrees bounded by the n_i and satisfying

$$A_0(z)P_0(z)+\cdots+A_k(z)P_k(z)=O(z^{n_0+\cdots+n_k+k}). \tag{1.2}$$

When $k=1$ and $A_1(z)=-1$, this is just the classical notion of a Padé approximation of a single power series. Other examples of Padé–Hermite approximation include the quadratic approximations of [23] and the D-Log approximations of [2]. Additional examples along with basic properties of Padé–Hermite approximants can be found in [3].

The concept of a Padé–Hermite approximant originated with ideas from the thesis of Padé [21] and some previous work of Hermite [12,13]. Hermite's ideas followed from his earlier study of a similar type of approximant for a vector of power series, the simultaneous Padé approximant (cf. [14]). This second type of approximant was used extensively by Hermite when he proved the transcedence of the number e. The general definition of both types of approximants, along with an extensive study of their properties is originally due to [18], with additional properties presented in [6,14].

A fundamental tool in the study of Padé approximants is the Padé table (cf. [9]). The Padé–Hermite table is a natural generalization of the Padé table. In [8] a number of relationships is discovered between neighboring entries in the table resulting in 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 in [22].

The resulting algorithms, however, cannot be applied to arbitrary power series. The algorithms of [8,22] require that the vector of power series be *perfect* (cf. [14]). Related to the concept of a Padé–Hermite approximation is a linear system of equations having a generalized Sylvester matrix as its coefficients. The condition of being perfect requires that the coefficient matrix, along with a specific set of submatrices, be nonsingular. This restriction 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 condition of being perfect. There are only a few known examples of perfect systems (cf. [6]).

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; the requirement of being perfect is not needed. Rather than using neighboring entries in the Padé–Hermite table, we introduce and use a new type of rational approximant, the weak Padé–Hermite approximant. This is a type of multi-dimensional rational approximant that is closely related to a simultaneous Padé approximant for the given set of power series (indeed, a weak Padé–Hermite approximant can easily be transformed into a specific set of simultaneous approximants). That simultaneous Padé approximants are closely related to Padé–Hermite approximants is well known; many important properties of their relationship are discussed in [19]. By viewing Padé–Hermite approximants as matrix polynomial rational approximants to a vector of power series, we are able to use many of the fundamental ideas of square matrix Padé approximants found in [16]. Our recurrence relation is a natural extension of the recurrence relation found in that paper.

In those cases where the coefficient matrix to the linear system is nonsingular, a particular Padé–Hermite approximant and weak Padé–Hermite approximant can be combined to make a Padé–Hermite system. Our recurrence relation results in an algorithm that calculates a desired approximant by computing these Padé–Hermite systems from one nonsingular 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 [5] and the scalar algorithm of [16]. 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 nonperfect case. In the perfect case, the algorithm is of the same complexity as the algorithms of [8,22].

2. Basic definitions

Let $n=(n_0,\ldots,n_k)$ be a vector of integers with $n_j \geqslant -1$ for all j. Following [22] we will denote

$$\|n\| = (n_0+1) + \cdots + (n_k+1) n_0 + \cdots + n_k + k + 1, \tag{2.1}$$

$$e=(1,\ldots,1) \quad \text{and} \quad e_0 = (1,0,\ldots,0). \tag{2.2}$$

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

$$A_i(z) = \sum_{j=0}^{\infty} a_{i,j} z^j, \quad i=0,\ldots,k, \tag{2.3}$$

be a set of $k+1$ formal power series with coefficients $a_{i,j}$ coming from a field F.

Definition 2.1. A $(k+1)\times 1$ vector of polynomials $P=(P_0,\ldots,P_k)^{\mathrm{T}}$ is defined to be a *Padé–Hermite form* (PHFo) of type n (with at least one $n_j \geqslant 0$) for the vector of power series $A=(A_0,\ldots,A_k)$ if

(I) P is nontrivial,

(II) $\partial(P_i) \leqslant n_i$, for $i=0,\ldots,k$,

(III) $$A(z)P(z) = (A_0(z),\ldots,A_k(z))(P_0(z),\ldots,P_k(z))^{\mathrm{T}} = z^{\|n\|-1} R(z), \tag{2.4}$$

where R is a power series (called the *residual* of type n).

For any positive integer λ let

$$T_{n,\lambda} = \begin{bmatrix} a_{0,0} & 0 & \cdots & 0 & & a_{k,0} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & & \vdots & \ddots & \ddots & \vdots \\ & & & 0 & & & & & \\ & & & a_{0,0} & \cdots & & & & 0 \\ & & & & & & & & a_{k,0} \\ \vdots & & & \vdots & & \vdots & & & \vdots \\ a_{0,\lambda-1} & \cdots & & a_{0,\lambda-n_0-1} & & a_{k,\lambda-1} & \cdots & & a_{k,\lambda-n_k-1} \end{bmatrix} \tag{2.5}$$

denote a generalized Sylvester matrix, and set

$$d_n = \det(T_{n,\|n\|}). \tag{2.6}$$

Theorem 2.2. *For any A and n (with at least one $n_i \geqslant 0$) there exists a PHFo of type n. If $d_n \neq 0$, then the corresponding PHFos are unique up to multiplication by a nonzero element of the field F. In addition, in this case the first term, $R(0)$, of the residual is nonzero.*

Proof. The problem of computing a PHFo of type n is equivalent to finding the solution X of the homogeneous system

$$T_{n,\|n\|-1}X = 0 \tag{2.7}$$

of $\|n\|-1$ linear equations and $\|n\|$ unknowns, which always has a nontrivial solution. Equivalently, we can solve the system

$$T_{n,\|n\|}X = [0,\ldots, 0, r_0]^{\mathrm{T}}, \tag{2.8}$$

with arbitrary $r_0 = R(0)$. When $d_n = \det(T_{n,\|n\|}) \neq 0$, a solution of (2.8) is uniquely defined (up to multiplication by an element of F). In addition, X is nontrivial if and only if $r_0 \neq 0$. □

Definition 2.3. A PHFo of type n with $R(0) = 1$ is called a *normed* PHFo.

Example 2.4. Let

$$A_0(z) = \cos(z) = 1 - \frac{z^2}{2} + \frac{z^4}{24} - \frac{z^6}{720} + \frac{z^8}{40320} - \frac{z^{10}}{3628800} + \cdots,$$

$$A_1(z) = \sin(z) = z - \frac{z^3}{6} + \frac{z^5}{120} - \frac{z^7}{5040} + \frac{z^9}{362880} - \frac{z^{11}}{39916800} + \cdots,$$

$$A_2(z) = \ln(1+z^2) + z^{15} = z^2 - \frac{z^4}{2} + \frac{z^6}{3} - \frac{z^8}{4} + \frac{z^{10}}{5} + \cdots.$$

The 9×9 matrix $T_{(4,3,-1),9}$ has nonzero determinant. Solving (2.8) with $r_0 = 1$ gives a normed PHFo of type $(4, 3, -1)$ by

$$P_0(z) = -1575z^2 + 105z^4, \qquad P_1(z) = 1575z - 630z^3, \qquad P_2(z) = 0.$$

Using (2.4) we obtain the first few terms of the residual as

$$R(z) = 1 - \frac{z^2}{18} + \frac{z^4}{792} - \frac{z^6}{61776} + \cdots.$$

3. Weak Padé–Hermite forms and fractions

For the remainder 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 $d_n \neq 0$). For clarity of presentation and without loss of generality, we assume this is true for $i = 0$, i.e., $a_{0,0} \neq 0$. The fact that $a_{0,0} \neq 0$ has some implications, which will prove useful later. Note that, if $n_i = -1$, for $i = 1,\ldots, k$, then $n_0 \geqslant 0$ and the PHFo is given trivially by

$$(P_0(z),\ldots, P_k(z)) = (z^{n_0}, 0,\ldots, 0). \tag{3.1}$$

As noted in Section 2, PHFos are equivalent to solutions of the linear system of equations (2.7) having $T_{n,\|n\|-1}$ as a coefficient matrix. Consider now the linear system that results from the deletion of the last $k-1$ rows of $T_{n,\|n\|-1}$, that is, $T_{n,\|n\|-k}$. The system (2.7) now is guaranteed to have, not one, but at least k linearly independent solutions. Each solution still satisfies (2.4) but only with a relaxed order condition. Thus, it is only a kind of PHFo, defined in this weakened sense. Such solutions are introduced in this paper primarily to facilitate the development in Section 5 of an algorithm for computing the genuine PHFos satisfying Definition 2.1.

Definition 3.1. The matrix $(P_1, \ldots, P_k)$ consisting of $(k+1) \times 1$ vector polynomials $P_j = (P_{0,j}, \ldots, P_{k,j})^T$ is called a *Weak Padé–Hermite Form* (WPHFo) for A of type n, where $n_i \geqslant 0$ for $0 \leqslant i \leqslant k$, if

(I) the columns of the $k \times k$ matrix $V = (P_{l,j})_{l=1,\ldots,k}^{j=1,\ldots,k}$ are linearly independent (with respect to coefficients from the field F),

(II) $\partial(P_{i,j}) \leqslant n_l$, for $1 \leqslant j \leqslant k$, $0 \leqslant l \leqslant k$,

(III) $A(P_1, \ldots, P_k) = z^{\|n\|-k}(R_1, \ldots, R_k)$, (3.2)

with the R_i a power series for all i.

The matrix polynomials $U = (P_{0,1}, \ldots, P_{0,k})$, V and $W = (R_1, \ldots, R_k)$ will be called the weak Padé–Hermite numerator, denominator and residual (all of type n), respectively. When $k = 1$, Definition 3.1 is the scalar definition of a Padé form (cf. [9]).

We may replace condition (I) of Definition 3.1 by the slightly weaker condition:

(I′) the columns of the $(k+1) \times k$ matrix $(P_1, \ldots, P_k) = (P_{l,j})_{l=0,\ldots,k}^{j=1,\ldots,k}$ are linearly independent.

To see that this is the case, suppose condition (I′) holds and let $\alpha = (\alpha_1, \ldots, \alpha_k)^T$, $\alpha_i \in F$, be such that

$$\alpha_1 P_{l,1} + \cdots + \alpha_k P_{l,k} = 0, \quad \text{for } l = 1, \ldots, k. \tag{3.3}$$

Then (3.3) and condition (II) of Definition 3.1 give

$$A(P_1, \ldots, P_k)\alpha = z^{\|n\|-k} R' = A_0 P', \tag{3.4}$$

with R' a power series and $P' = \alpha_1 P_{0,1} + \cdots + \alpha_k P_{0,k}$. Multiplying both sides with the reciprocal power series of A_0 (using the fact that $A_0(0) = a_{0,0} \neq 0$) gives

$$P' = z^{\|n\|-k} \hat{R}, \tag{3.5}$$

with $\hat{R}$ a power series. By condition (II), $\partial(P') \leqslant n_0$ which in (3.5) is less than the order condition $\|n\| - k = n_0 + \cdots + n_k + 1$. Hence

$$P' = \alpha_1 P_{0,1} + \cdots + \alpha_k P_{0,k} = 0, \tag{3.6}$$

also holds and condition (I′) then implies that $\alpha_1 = \cdots = \alpha_k = 0$.

As a consequence of the above observation, computing a WPHFo of type n is equivalent to finding k linearly independent solutions Y of the homogeneous linear system

$$T_{n,\|n\|-k} Y = 0 \tag{3.7}$$

of $\|n\| - k$ linear equations and $\|n\|$ unknowns. This yields the following theorem.

Theorem 3.2 (Existence of WPHFos). *For any A and n with $A_0(0) \neq 0$ and $n_i \geqslant 0$ for all i, there exists a WPHFo of type n.*

Definition 3.3. A WPHFo of type n with a nonsingular matrix $V(0)=(P_{j,l}(0))_{l=1,\dots,k}^{j=1,\dots,k}$ is called a *Weak Padé–Hermite Fraction* (WPHFr). If $V(0)$ equals the $k\times k$ identity matrix, then we call it a *normed* WPHFr.

Note that any WPHFr $(P_1,\dots,P_k)$ can be made into a normed WPHFr by multiplication on the right by $V(0)^{-1}$. Also note that, when $k=1$, a WPHFr is the same as a scalar Padé fraction (cf. [9]).

A WPHFr can be interpreted as providing a set of simultaneous Padé approximants for the quotient power series $A_i(z)/A_0(z)$ (cf. [7]). We may write (3.2) as

$$A_0(z)U(z)+(A_1(z),\dots,A_k(z))V(z)=z^{\|n\|-k}(R_1(z),\dots,R_k(z)), \tag{3.8}$$

that is,

$$A_0(z)U(z)+(A_1(z),\dots,A_k(z))V(z)\approx 0. \tag{3.9}$$

Since $V(0)$ is nonsingular, the inverse of the $k\times k$ matrix polynomial $V(z)$ can be determined as a matrix power series. Thus, we obtain

$$\left(\frac{A_1(z)}{A_0(z)},\dots,\frac{A_k(z)}{A_0(z)}\right)\approx -U(z)V(z)^{-1}. \tag{3.10}$$

Since

$$U(z)V(z)^{-1}=U(z)\frac{\operatorname{adj}(V(z))}{\det(V(z))}, \tag{3.11}$$

(3.10) and (3.11) give a simultaneous rational approximation for each power series

$$\frac{A_i(z)}{A_0(z)}\approx\frac{N_i(z)}{D(z)},\quad i=1,\dots,k. \tag{3.12}$$

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 $(D(z),N_1(z),\dots,N_k(z))$ form a set of simultaneous Padé approximants to the power series $A_1(z)/A_0(z),\dots,A_k(z)/A_0(z)$ of type n. This is also called the German polynomial approximation problem of type n for the power series $A_0(z),\dots,A_k(z)$ or as directed vector Padé approximants for the vector of power series $(A_0(z),\dots,A_k(z))$ in the unit coordinate directions (cf. [10]). In [4] also definitions are given of matrix Padé approximants similar to WPHFos and WPHFrs for use in computing minimal partial realizations.

By deleting the first column of the lth block of $T_{n,\|n\|-k}$ for $l=1,\dots,k$, we obtain the square matrix

$$T_n^*=\begin{bmatrix} a_{0,0} & 0 & \cdots & 0 & 0 & \cdots & 0 & & 0 & \cdots & 0 \\ \vdots & a_{0,0} & \ddots & \vdots & a_{1,0} & \ddots & \vdots & & a_{k,0} & \ddots & \vdots \\ & \vdots & \ddots & & & \ddots & 0 & & & \ddots & 0 \\ & & & 0 & \vdots & & a_{1,0} & \cdots & \vdots & & a_{k,0} \\ & & & a_{0,0} & & & \vdots & & & & \vdots \\ & & & \vdots & & & & & & & \\ a_{0,\eta} & a_{0,\eta-1} & & a_{0,\eta-n_0} & a_{1,\eta-1} & & a_{1,\eta-n_1} & \cdots & a_{k,\eta-1} & & a_{k,\eta-n_k} \end{bmatrix}, \tag{3.13}$$

with

$$\det(T_n^*) = a_{0,0} \det(T_{n-e,\|n-e\|}) = a_{0,0} d_{(n-e)}, \tag{3.14}$$

where e is given by (2.2) and $\eta = \|n\| - k - 1$.

Theorem 3.4. *If $d_{(n-e)} \neq 0$, then every WPHFo of type n is a WPHFr. In addition, any WPHFr of type n is unique up to multiplication on the right by a nonsingular $k \times k$ matrix with coefficients from the field F. In particular, there exists one and only one normed WPHFr of type n.*

Proof. By assumption we know that the $(\|n\| - k) \times (\|n\| - k)$ square submatrix T_n^* of $T_{n,\|n\|-k}$ has full rank, i.e., $\operatorname{rank}(T_{n,\|n\|-k}) = \|n\| - k$. Therefore the system of equations (3.7) has exactly k linearly independent solutions and a WPHFo is unique up to a multiplication on the right by a nonsingular $k \times k$ matrix.

The first column of the jth block of $T_{n,\|n\|-k}$, for $j = 1, \ldots, k$, corresponds to the unknowns $P_{l,j}(0)$, for $l = 1, \ldots, k$. Hence in order to obtain k linearly independent solutions of (3.7), we first have to choose k linearly independent parameter vectors $(P_{l,j}(0))_{j=1,\ldots,k}$ and then compute the matrix Y of the other unknowns by

$$T_n^* Y = -\begin{bmatrix} a_{1,0} & \cdots & a_{k,0} \\ \vdots & & \vdots \\ a_{1,\eta} & \cdots & a_{k,\eta} \end{bmatrix} V(0). \tag{3.15}$$

Thus each WPHFo is a WPHFr. □

Example 3.5. Let $A_0(z)$, $A_1(z)$, $A_2(z)$ be as in Example 2.4. We will verify Theorem 3.4 in the case $n = (4, 3, -1)$. The normed PHFo of type n is given by Example 2.4. After some computations we obtain for the normed WHPFr of type (5, 4, 0):

$$U(z) = \left[-z + \frac{z^3}{9} - \frac{z^5}{945},\ -\frac{2773}{8}z^2 + \frac{287}{12}z^4 \right],$$

$$V(z) = \begin{bmatrix} 1 - \frac{4}{9}z^2 + \frac{z^4}{63} & \frac{2765}{8}z - \frac{1113}{8}z^3 \\ 0 & 1 \end{bmatrix},$$

with the residual

$$W(z) = \left[\frac{z}{9823275} - \frac{z^3}{255405150} + \cdots,\ \frac{13553}{72576} - \frac{2213653}{13305600}z^2 + \cdots \right].$$

If desired, the matrix polynomials $U(z)$ and $V(z)$ can be transformed via (3.10) and (3.11) into simultaneous Padé approximants of type (5, 4, 0) for $A_1(z)/A_0(z)$ and $A_2(z)/A_0(z)$ given by

$$\frac{945z - 105z^3 + z^5}{945 - 420z^2 + 15z^4} \quad \text{and} \quad \frac{7560z^2 - 3360z^4 + 2325z^6 - 1757z^8}{7560 - 3360z^2 + 120z^4},$$

respectively.

Central to our results is the fact that Theorems 2.2 and 3.4 actually classify when the determinant of the generalized Sylvester matrix is nonzero. The following definition is helpful in arriving at these results.

Definition 3.6. Let $P = (P_0, \ldots, P_k)$ be a $(k+1) \times (k+1)$ matrix of polynomials, $n = (n_0, \ldots, n_k)$ a vector of integers with $n_i > -1$ for at least one i. Then P is called a *Padé–Hermite System* (PHS) of type n, if

(I) P_0 is a normed PHFo of type n,

(II) $(P_1, \ldots, P_k)$ is a normed WPHFr of type $n + e$.

Theorem 3.7 *There exists a PHS of type n if and only if $d_n \neq 0$.*

Proof. Theorems 2.2 and 3.4 show that $d_n \neq 0$ is sufficient for the existence of a PHS of type n. Suppose that there is a PHS $(P_0, \ldots, P_k)$ of type n. Set $\lambda = \| n \|$. Then equations (2.8) and (3.15) imply the existence of a $\lambda \times 1$ solution vector X of

$$T_{n,\lambda} X = [0, \ldots, 0, 1]^{\mathrm{T}}, \tag{3.16}$$

and a $(\lambda + 1) \times k$ solution matrix Y of

$$T^*_{n+e} Y = -\begin{bmatrix} a_{1,0} & \cdots & a_{k,0} \\ \vdots & & \vdots \\ a_{1,\lambda} & \cdots & a_{k,\lambda} \end{bmatrix}. \tag{3.17}$$

If we assume that $d_n = 0$, then $T_{n,\lambda}$ is singular. Thus there exists a nontrivial solution to the homogeneous system of equations

$$(s_1, \ldots, s_\lambda) T_{n,\lambda} = 0. \tag{3.18}$$

Equation (3.16) together with (3.18) then yield

$$s_\lambda = (s_1, \ldots, s_\lambda) T_{n,\lambda} X = 0. \tag{3.19}$$

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

$$(\tau, s_1, \ldots, s_\lambda)(a_{0,0}, \ldots, a_{0,\lambda})^{\mathrm{T}} = 0. \tag{3.20}$$

Then,

$$(\tau, s_1, \ldots, s_\lambda) T^*_{n+e} = 0. \tag{3.21}$$

Equations (3.17) and (3.21) yield

$$(\tau, s_1, \ldots, s_\lambda) \begin{bmatrix} a_{1,0} & \cdots & a_{k,0} \\ \vdots & & \vdots \\ a_{1,\lambda} & \cdots & a_{k,\lambda} \end{bmatrix} = 0. \tag{3.22}$$

Therefore, from (3.18), (3.20) and (3.22) we get

$$(\tau, s_1, \ldots, s_\lambda) T_{n,\lambda+1} = 0. \tag{3.23}$$

Equation (3.23) coupled with $s_\lambda = 0$ implies that

$$(\tau, s_1, \ldots, s_{\lambda-1}) T_{n,\lambda} = 0. \tag{3.24}$$

An induction argument can then be used to show that

$$s_\lambda = s_{\lambda-1} = \cdots = s_1 = 0, \tag{3.25}$$

which contradicts the assumption that there is a nontrivial solution to (3.18). Hence, $T_{n,\lambda}$ is nonsingular and $d_n \neq 0$. □

Remark 3.8. When $k=1$, Theorem 3.7 was proved in [17]. Note that, as a consequence of Theorem 3.7, there exists at most one PHS of type n.

Remark 3.9. A vector of power series A is said to be *perfect* if $d_n \neq 0$ for all integer vectors n (cf. [14]). Theorem 3.7 implies that perfect vectors of power series are precisely those having a PHS of type n for all integer vectors n.

Remark 3.10. Notice that the proof of Theorem 3.7 gives necessary and sufficient conditions for the nonsingularity of a generalized Sylvester matrix. Indeed, if

$$S = \begin{bmatrix} s_{0,0} & \cdots & s_{0,m_0} & \cdots & s_{k,0} & \cdots & s_{k,m_k} \\ \vdots & & \vdots & & \vdots & & \vdots \\ s_{0,N} & \cdots & s_{0,N+m_0} & \cdots & s_{k,N} & \cdots & s_{k,N+m_k} \end{bmatrix}, \tag{3.26}$$

with $N = m_0 + \cdots + m_k + k$, then, using arguments similar to those used in the proof of Theorem 3.7, it can be seen that S is nonsingular if and only if there exist solutions to the equations

$$S\left(x_0^{(i)},\ldots, x_N^{(i)}\right)^{\mathrm{T}} = -\left(s_{i,m_i+1},\ldots, s_{i,N+m_i+1}\right)^{\mathrm{T}}, \quad \text{for } i \in 0,\ldots, k, \tag{3.27}$$

$$S(y_0,\ldots, y_N)^{\mathrm{T}} = (0,\ldots, 0, 1)^{\mathrm{T}}. \tag{3.28}$$

Similar results can be found in [11].

4. A recurrence relation for Padé–Hermite approximants

Given a vector of power series (2.3) and a vector of integers n, a corresponding PHFo can be determined by solving (2.7) via a method such as Gaussian elimination. 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 WPHFos via the solution to the system (3.7). However, such calculations do not take into account the special structure of the coefficient matrices of these sytems. 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 matrices (2.6) and (3.11), and at the same time it will not require any restrictions on the input.

Given a vector of power series (2.3), along with a vector $n = (n_0,\ldots, n_k)$ of nonnegative integers, we will permute the components so that

$$A_0(0) \neq 0, \qquad n_0 = \max\{n_i: A_i(0) \neq 0\} \quad \text{and} \quad n_1 \geqslant \cdots \geqslant n_k. \tag{4.1}$$

Note, that if $A_i(0) = 0$ for all $0 \leqslant i \leqslant k$, then it is only necessary to remove the largest factor z^β

from all the power series. Any PHFo or WPHFo of type n for $(z^{-\beta}A_0(z),\ldots, z^{-\beta}A_k(z))$ is then also a PHFo or WPHFo, respectively, of the same type for $(A_0(z),\ldots, A_k(z))$.

Set $M=\max(n_0+1,\ldots, n_k+1)$, and define integer vectors $n^{(i)}=(n_0^{(i)},\ldots, n_k^{(i)})$ for $0\leqslant i\leqslant M$ by

$$n_j^{(i)}=\max(-1, n_j-M+i),\quad \text{for } j=0,\ldots, k. \tag{4.2}$$

Then the sequence $\{n^{(i)}\}_{i=0,1,\ldots}$ lies on a piecewise linear path with $n_j^{(i+1)}\geqslant n_j^{(i)}$ for each i, j and

$$n^{(0)}=(-1,\ldots, -1)\quad\text{and}\quad n^{(M)}=(n_0,\ldots, n_k)=n. \tag{4.3}$$

The sequence $\{n^{(i)}\}$ generates a subsequence $\{m^{(i)}\}$ called the *sequence of nonsingular points*. This sequence is defined by $m^{(i)}=n^{(\sigma_i)}$ where

$$\sigma_i=\begin{cases}0 & i=0,\\ \min(\sigma>\sigma_{i-1}\colon d_{n^{(\sigma)}}\neq 0), & i\geqslant 1.\end{cases} \tag{4.4}$$

Observe that the ordering (4.1) implies $m_0^{(i)}\geqslant 0$ for all $i\geqslant 1$; and therefore, for all $\sigma>\sigma_i>0$ it is true that

$$\sigma-\sigma_i=n_0^{(\sigma)}-m_0^{(i)}. \tag{4.5}$$

Consequently, there exists $1\leqslant a_1\leqslant a_2\leqslant\cdots$ such that for all $i\geqslant 1$ we have

$$m_j^{(i)}=\begin{cases}\geqslant 0, & \text{for } j=0,\ldots, a_i-1,\\ \geqslant -1, & \text{for } j=a_i,\ldots, k.\end{cases} \tag{4.6}$$

Example 4.1. Let $A_0(z)$, $A_1(z)$ and $A_2(z)$ be given by Example 2.4 with $n=(7, 6, 1)$. Then $M=8$ and calculating the required determinants gives the first four nonsingular points as

$$m^{(1)}=n^{(1)}=(0, -1, -1),\qquad m^{(2)}=n^{(3)}=(2, 1, -1),$$
$$m^{(3)}=n^{(5)}=(4, 3, -1),\qquad m^{(4)}=n^{(8)}=(7, 6, 1).$$

The quantities a_i are then given as $a_1=1$, $a_2=2$, $a_3=2$, $a_4=3$.

For $i\geqslant 1$, let $P^{(i)}=(P_0^{(i)},\ldots, P_k^{(i)})$ be the uniquely defined PHS of type $m^{(i)}$ (cf. Definition 3.6 and Theorem 3.7), with $R^{(i)}=(R_0^{(i)},\ldots, R^{(i)})$ its residual vector. We will uses the partitions

$$P^{(i)}=\begin{bmatrix}P_{0,0}^{(i)} & U^{(i)}\\ Q^{(i)} & V^{(i)}\end{bmatrix},\qquad R^{(i)}=\left[R_0^{(i)}|W^{(i)}\right], \tag{4.7}$$

where by definition

$$(A_0(z),\ldots, A_k(z))P^{(i)}(z)=z^{\|m^{(i)}\|-1}\left(R_0^{(i)}(z)|z^2W^{(i)}(z)\right), \tag{4.8}$$

and $R_0^{(i)}(0)=1$, $V^{(i)}(0)=I_k$.

The algorithm described in Section 5 for constructing a PHFo of type n for a vector of power series A involves the computation of all PHFos and WPHFrs up to the point n. Theorem 4.2 gives a relationship of the $(i+1)$st terms of the sequences with the ith terms, providing an effective mechanism for computing the sequences.

Theorem 4.2. *For $i \geqslant 1$, $\sigma > \sigma_i$, let $\nu = n^{(\sigma)} - m^{(i)} - e_0 - e$. Then $n^{(\sigma)}$ is a nonsingular point for $A = (A_0, \dots, A_k)$ if and only if ν is a nonsingular point for $R^{(i)} = (R_0^{(i)}, \dots, R_k^{(i)})$. Furthermore, we have the recurrence relations*

$$P^{(i+1)} = \left(z^2 P_0^{(i)}, P_1^{(i)}, \dots, P_k^{(i)}\right)(P_0', \dots, P_k') \quad \text{and} \quad R^{(i+1)} = R', \tag{4.9}$$

where $P' = (P_0', \dots, P_k')$ is the PHS of type $(m^{(i+1)} - m^{(i)} - e_0 - e)$ for the system $R^{(i)}$ and R' is its residual.

Proof. Let P_0' and $(P_1', \dots, P_k')$ be a PHFo of type ν and a WPHFo of type $\nu + e$, respectively, for $(R_0^{(i)}, \dots, R_k^{(i)})$. Let R_0' and $(R_1', \dots, R_k')$ be the corresponding residuals. Further, set

$$P^* = \left(z^2 P_0^{(i)}, P_1^{(i)}, \dots, P_k^{(i)}\right)(P_0', \dots, P_k') \quad \text{and} \quad R^* = (R_0', \dots, R_k'). \tag{4.10}$$

With $s = n_0^{(\sigma)} - m_0^{(i)} \geqslant 1$, $a = a_i$, we get

$$\nu = (\nu_0, \dots, \nu_k), \quad \text{with} \quad \nu_j = \begin{cases} s-2, & \text{for } j = 0, \\ s-1, & \text{for } 1 \leqslant j < a, \\ n_j^{(\sigma)}, & \text{for } j \leqslant a \leqslant k. \end{cases} \tag{4.11}$$

Notice that $\nu_j \leqslant s - 1$ for all j. For $0 \leqslant j \leqslant k$, $a \leqslant l \leqslant k$, we obtain $\partial(P_{l,j}^{(i)}) \leqslant 0$ and, because $V^{(i)}(0)$ is the identity,

$$\partial\left(P_{l,j}^{(i)}\right) \leqslant -1 + \delta_{j,l}. \tag{4.12}$$

Hence for the degrees of the entries of the following matrices we get on an element by element basis:

$$\partial\left(z^2 P_0^{(i)}, P_1^{(i)}, \dots, P_k^{(i)}\right) \leqslant \begin{bmatrix} m_0^{(i)}+2 & m_0^{(i)}+1 & \cdots & & m_0^{(i)}+1 \\ \vdots & \vdots & & & \vdots \\ m_{a-1}^{(i)}+2 & m_{a-1}^{(i)}+1 & \cdots & & m_{a-1}^{(i)}+1 \\ -1 & -1 & 0 & -1 & -1 \\ \vdots & \vdots & -1 & & \vdots \\ & & & & -1 \\ -1 & -1 & -1 & -1 & 0 \end{bmatrix}, \tag{4.13}$$

where the last $k - a + 1$ columns have the 0 entries. Also,

$$\partial(P_0', \dots, P_k') \leqslant \begin{bmatrix} s-2 & s-1 & \cdots & s-1 \\ s-1 & s & & s \\ \vdots & \vdots & & \vdots \\ s-1 & s & \cdots & s \\ n_a^{(\sigma)} & n_a^{(\sigma)}+1 & \cdots & n_a^{(\sigma)}+1 \\ \vdots & \vdots & & \vdots \\ n_k^{(\sigma)} & n_k^{(\sigma)}+1 & \cdots & n_k^{(\sigma)}+1 \end{bmatrix} \tag{4.14}$$

(where the first a rows have the s terms), so that

$$\partial(P_0^*,\ldots,P_k^*) \leqslant \begin{bmatrix} m_0^{(i)}+s & m_0^{(i)}+s+1 & \cdots & m_0^{(i)}+s+1 \\ \vdots & \vdots & & \vdots \\ m_{a-1}^{(i)}+s & m_{a-1}^{(i)}+s+1 & \cdots & m_{a-1}^{(i)}+s+1 \\ n_a^{(\sigma)} & n_a^{(\sigma)}+1 & \cdots & n_a^{(\sigma)}+1 \\ \vdots & \vdots & & \vdots \\ n_k^{(\sigma)} & n_k^{(\sigma)}+1 & \cdots & n_k^{(\sigma)}+1 \end{bmatrix}$$

$$= \begin{bmatrix} n_0^{(\sigma)} & n_0^{(\sigma)}+1 & \cdots & n_0^{(\sigma)}+1 \\ \vdots & \vdots & & \vdots \\ n_k^{(\sigma)} & n_k^{(\sigma)}+1 & \cdots & n_k^{(\sigma)}+1 \end{bmatrix}. \tag{4.15}$$

Also,

$$\begin{aligned}(A_0(z),\ldots,A_k(z))P^*(z) &= z^{\|m^{(i)}\|-1}z^2(R_0^{(i)},\ldots,R_k^{(i)})P'(z) \\ &= z^{\|m^{(i)}\|-1}z^{\|\nu\|-1}(R_0'(z),\ z^2R_1'(z),\ldots,\ z^2R_k'(z)) \\ &= z^{\|n^{(\sigma)}\|-1}(R_0^*(z),\ z^2R_1^*(z),\ldots,\ z^2R_k^*(z)),\end{aligned} \tag{4.16}$$

so the R_j^* are power series with $R_j^* = R_j'$.

In addition, partitioning P' in a similar manner as (4.7), equation (4.10) implies

$$V^*(z) = z^2Q^{(i)}(z)U'(z) + V^{(i)}(z)V'(z), \tag{4.17}$$

so

$$V^*(0) = V^{(i)}(0)V'(0) = V'(0). \tag{4.18}$$

Hence $R_0'(0) = 1$ if and only if $R_0^*(0) = 1$ and $V'(0) = I_k$ if and only if $V^*(0) = I_k$. Therefore, P' is a PHS of type ν for $(R_0^{(i)},\ldots,R_k^{(i)})$ if and only if P^* is a PHS of type $n^{(\sigma)}$ for $(A_0,\ldots,A_k)$. □

Remark 4.3. Theorem 4.2 reduces the problem of determining a PHS of type n to two "smaller" problems: determine a PHS of type $m^{(i)}$ and then determine a PHS of type $m^{(i+1)} - m^{(i)} - e - e_0$. 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 the multiplication in (4.9). This overhead cost is generally an order of magnitude less than the cost of simply solving the linear systems (2.7) or (3.7).

Remark 4.4. Theorem 2.2 implies that PHFos at nonsingular points of the Padé–Hermite table are unique up to multiplication by a scalar. Thus a normed PHFo represents a type of canonical form at such points. It is a natural question to ask for canonical forms at singular points in the Padé–Hermite table. All solutions at a given point could then be expressed in terms of the canonical one (see [1] or [4] where this is accomplished for matrix Padé approximants). Theorem 4.2 implies that there is a one-to-one correspondence between PHFos (and WPHFos) of A along a piecewise linear path from one nonsingular point to the next, and PHFos (and WPHFos) of a residual along a correspondence piecewise linear path to a first nonsingular point. Thus we only need to determine the structure of the Padé–Hermite table (along with any canonical forms) along the first singular path of any vector of power series. This will then give the structure of the Padé–Hermite table for arbitrary vectors of power series. In the $k=1$ case this provides a simple mechanism for exhibiting the block structure of the classical Padé table.

Example 4.5. Let $A_0(z)$, $A_1(z)$ and $A_2(z)$ be as in Example 2.4. Then $(4, 3, -1)$ is a nonsingular point. By Theorem 4.2, to determine the next nonsingular point we need only determine the first nonsingular point for $(R_0^{(3)}(z), R_1^{(3)}(z), R_2^{(3)}(z))$. In this case

$$\nu = \left(n_0^{(\sigma)} - 6,\ n_1^{(\sigma)} - 4,\ n_2^{(\sigma)}\right) = (\sigma - 7,\ \sigma - 6,\ \sigma - 7), \quad \text{for } 6 \leqslant \sigma \leqslant 8,$$

and the first nonsingular point occurs when $\nu = (1, 2, 1)$. If we are only interested in the PHFo of type $(7, 6, 1)$ for A, then we need only determine the PHFo of type $(1, 2, 1)$ for $R^{(3)}$, which in this case is

$$P'_{0,0}(z) = \frac{1\,058\,838\,979}{2\,903\,040} z, \qquad P'_{1,0}(z) = -\frac{38\,320\,755\,508\,035}{1024} + \frac{328\,312\,406\,639}{512} z^2,$$

$$P'_{2,0}(z) = z.$$

This represents the first column of the PHS of type $(1, 2, 1)$ for the residual power series. Using (4.9), the first column of the PHS of type $(7, 6, 1)$ for A is

$$P_{0,0}^{(4)}(z) = \frac{38\,320\,755\,508\,035}{1024} z - \frac{4\,914\,486\,891\,337}{1024} z^3 + \frac{53\,682\,800\,837}{512} z^5 - \frac{1\,426\,272\,217}{5120} z^7,$$

$$P_{1,0}^{(4)}(z) = -\frac{38\,320\,755\,508\,035}{1024} + \frac{8\,844\,036\,029\,829}{512} z^2 - \frac{446\,977\,776\,911}{512} z^4 + \frac{5\,973\,822\,233}{768} z^6,$$

$$P_{2,0}^{(4)}(z) = z.$$

This gives the PHFo of type $(7, 6, 1)$ for A. Using (2.4) gives the residual as

$$R_0^{(4)}(z) = 1 - \frac{49\,762\,803\,403}{411\,675\,264\,000} z + \frac{750\,531\,581\,615\,899}{6\,758\,061\,133\,824\,000} z^3 + \cdots.$$

Example 4.6. In the special case when $k=1$, a WHPFr is the same as a Padé fraction. In this case (4.8) is given by

$$A_0(z) U^{(i)}(z) + A_1(z) V^{(i)}(z) = z^{m_0^{(i)} + m_1^{(i)} + 3} W^{(i)}(z), \tag{4.19}$$

and $(U^{(i)}(z), V^{(i)}(z))$ is a Padé fraction of type $(m_0^{(i)} + 1,\ m_1^{(i)} + 1)$ for $(A_0(z), A_1(z))$. If

$$W^{(i)}(z) = z^{\lambda_i} \hat{W}(z), \tag{4.20}$$

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

$$P_{0,0}^{(i+1)}(z) = z^{\lambda_i}\hat{w}_0^{-1}U^{(i)}(z), \qquad Q^{(i+1)}(z) = z^{\lambda_i}\hat{w}_0^{-1}V^{(i)}(z), \tag{4.21}$$

with residual

$$R_0^{(i+1)}(z) = \hat{w}_0^{-1}\hat{W}^{(i)}(z). \tag{4.22}$$

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

When $k = 1$, the Extended Euclidean Algorithm for computing polynomial GCDs is closely related to the calculation of Padé approximants (cf. [5,20]). When the input power series $A_0(z)$ and $A_1(z)$ are polynomials of degree m and n, respectively, then reversing the order of the coefficients in (4.19) and using (4.21) and (4.22) gives

$$A_0^*(z)P_{0,0}^{*(i)}(z) + A_1^*(z)Q^{*(i)}(z) = R_0^{*(i)}(z). \tag{4.23}$$

Here $A_0^*(z) = z^m A_0(z^{-1}),\ldots,$ etc. Equation (4.23) 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) for $(A_0(z), A_1(z))$, the reversed residual $R_0^{*(i)}(z)$ is the ith term of the remainder sequence calculated in the EEA, while $(P_{0,0}^{*(i)}(z), Q^{*(i)}(z))$ is the ith term of the cofactor sequence calculated in the EEA. In this way our recurrence relation, and the subsequent algorithm presented in the next section can be viewed as a generalization of Euclidean algorithm to computing Padé–Hermite approximants for vectors of power series. Other generalizations of Euclid's algorithm include that given in [1] (applied to the computation of minimal partial realizations) and [4] (applied to the problem of computing minimal matrix Padé approximants).

5. The algorithm

Given a vector of nonnegative integers $n = (n_0,\ldots, n_k)$, the algorithm PADE_HERMITE below makes use of Theorem 4.2 to compute the sequence $\{P^{(i)}\}$ of PHS for a given vector of power series $A = (A_0,\ldots, A_k)$. Thus, intermediate results available from PADE_HERMITE include those PHFos for A at all nonsingular points $m^{(i)}$, $i = 1, 2,\ldots, l-1$, smaller than n on the piecewise linear path defined by the sequence $\{n^{(i)}\}$, together with those WPHFr at the succeeding points. The output gives results associated with the final point $m^{(l)}$. If this final point is a nonsingular point, then the output $(P_0^{(l)},\ldots, P_k^{(l)})$ is a PHS of type n. If n is a singular point, then the output is $(P_0^{(l)},\ldots, P_k^{(l)})$, where the polynomial vector $P_0^{(l)}$ is a PHFo of type n, while $(P_1^{(l)},\ldots, P_k^{(l)})$ is a WPHFo at the successor point.

The algorithm is presented in two parts. The first, INITIAL_PH, takes as its input a vector of power series, A, with $A_0(0) \neq 0$ and an integer vector n with $n_1 \geqslant \cdots \geqslant n_k$. The procedure returns the PHS at the first nonsingular point, if such a point exists. Otherwise a PHFo and a WPHFo are calculated and returned (note that in this case a number of possibilities may exist for such a PHFo and WPHFo). The chosen PHFo and WPHFo are arranged into a $(k+1) \times (k+1)$ polynomial matrix with the PHFo the first column and the WPHFo the last k columns.

The main algorithm PADE_HERMITE calls INITIAL_PH to iteratively construct PHSs for the residuals $R^{(i)}(z)$. The PHSs $P^{(i)}$ for A are computed using the results of Theorem 4.2. In the case where INITIAL_PH does not return a PHS, then PADE_HERMITE returns a PHFo and WPHFo of correct type.

INITIAL_PH(A, n)

(I.1) $d \leftarrow 0$; $M \leftarrow \max(n_0, n_1)+1$; $\sigma \leftarrow 0$
(I.2) do while $\sigma < M$ and $d=0$
(I.3) $\sigma \leftarrow \sigma + 1$
(I.4) $n^{(\sigma)} \leftarrow \max(-1, n_j - M + \sigma)$, $j=0,\ldots, k$,
(I.5) compute $d \leftarrow \det(T_{n^{(\sigma)}, \| n^{(\sigma)} \|})$, using Gaussian elimination
end while
(I.6) if $d \neq 0$, then solve (3.16) and (3.17) for P, the PHS of type $n^{(\sigma)}$ for A; else solve the corresponding homogeneous equations; arrange the solutions into a $(k+1)\times(k+1)$ matrix P and set $\sigma \leftarrow M+1$
(I.7) return (σ, P)

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 nonnegative entries (otherwise one calls PADE_HERMITE with a smaller value of k).

PADE_HERMITE(A, n)

(PH.1) find the largest β such that $A_i(z) = z^{\beta}\hat{A}_i(z)$ are still power series; set $A_i(z) = z^{-\beta}A_i(z)$; reorder the power series according to (4.1)
(PH.2) $M \leftarrow \max(n_0, n_1)+1$
(PH.3) $(s_0, P^{(1)}) \leftarrow$ INITIAL_PH(A, n)
(PH.4) $\sigma \leftarrow s_0$; $m_j^{(1)} \leftarrow \max(-1, n_j - M + \sigma)$, $j=0,\ldots, k$; $i \leftarrow 1$
(PH.5) while $\sigma \leqslant M$ do
(PH.6) determine $R^{(i)}$ using (4.8); $\nu \leftarrow n - m^{(i)} - e - e_0$
(PH.7) $(s_{i-1}, P') \leftarrow$ INITIAL_PH($R^{(i)}$, ν)
(PH.8) $[P_0^{(i+1)},\ldots, P_k^{(i+1)}] \leftarrow [z^2P_0^{(i)}, P_1^{(i)},\ldots, P_k^{(i+1)}][P_0',\ldots, P_k']$
(PH.9) $\sigma \leftarrow \sigma + s_{i-1}$; $m_j^{(i+1)} \leftarrow \max(-1, n_j - M + \sigma)$, $j=0,\ldots, k$; $i \leftarrow i+1$
end while
(PH.10) return $(\sigma, [P_0^{(i)},\ldots, P_k^{(i)}])$

Example 5.1. Let $A_0(z)$, $A_1(z)$ and $A_2(z)$ be as in Example 2.4. Then the first nonsingular point is at $(0, -1, -1)$. The values of the PHFo of type $(0, -1, -1)$ and the WPHFr of type $(1, 0, 0)$ are determined by solving the linear system of equations (3.16) and (3.17) to give the PHS of type $(0, -1, -1)$ as

$$P^{(1)} = \begin{bmatrix} 1 & -z & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The residuals (calculated using (4.8)) are given by

$$R^{(1)} = \left[1 - \tfrac{1}{2}z^2 + \cdots,\ \tfrac{1}{3}z - \tfrac{1}{30}z^3 + \cdots,\ 1 - \tfrac{1}{2}z^2 + \cdots\right].$$

The first nonsingular point for the residuals amongst the sequence $(-1, 0, -1)$, $(0, 1, -1)$,

$(1, 2, -1), \ldots$ is found at location $(0, 1, -1)$. The corresponding PHS of type $(0, 1, -1)$ is given by

$$P' = \begin{bmatrix} 0 & -\frac{1}{3}z & -1 \\ 3z & 1 - \frac{2}{5}z^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Combining these via step (PH.8) gives the PHS of type $(2, 1, -1)$ for A as

$$P^{(2)} = \begin{bmatrix} -3z^2 & -z + \frac{1}{15}z^3 & -z^2 \\ 3z & 1 - \frac{2}{5}z^2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

along with residuals

$$R^{(2)} = \left[1 - \frac{z^2}{10} + \cdots, \frac{z}{1575} - \frac{z^3}{28350} + \cdots, \frac{7}{24} - \frac{179}{720}z^2 + \cdots\right].$$

Continuing with the next step, the first nonsingular point for the residual amongst the sequence $\{(-1, 0, -1), (0, 1, -1), (1, 2, -1), \ldots\}$ occurs at location $(0, 1, -1)$. The corresponding PHS is given by

$$P' = \begin{bmatrix} 0 & -\frac{z}{1575} & -\frac{7}{24} \\ 1575z & 1 - \frac{2}{45}z^2 & \frac{2765}{8}z \\ 0 & 0 & 1 \end{bmatrix}.$$

Combining these via step (PH.8) and using the partition (4.8) gives the PHFo and WPHFrs previously calculated in Examples 2.4 and 3.5, respectively. The final step (when calculating only the PHFo of type $(7, 6, 1)$) has been done previously in Example 4.5.

6. Complexity of the 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 is used in step (I.5) to obtain a triangular factorization of $T_{n^{(\sigma)}, \|n^{(\sigma)}\|}$. Assuming that the elimination is performed by applying bordering techniques (as σ increases), step (I.5) requires approximately $\frac{1}{3}(\|n^{(\sigma)}\|)^3$ multiplications in F, where $n^{(\sigma)}$ is the integer vector attained upon exit from the while loop (I.2). In the case where $d \neq 0$, the solution of the equations P resulting in the PHS of type $n^{(\sigma)}$ can then be obtained by forward and backward substitution requiring approximately $(k+1)(\|n^{(\sigma)}\|)^2$ multiplications in total. Since at the ith invocation of INITIAL_PH, $n^{(\sigma)} = \nu_i$, where $\nu_i = \|m^{(i+1)}\| - \|m^{(i)}\| - 1$, the total cost of this invocation at the ith iteration of PADE_HERMITE is approximately

$$\tfrac{1}{3}\nu_i^3 + \nu_i^2(k+1). \tag{6.1}$$

If we set $\mu_i = \| m^{(i)} \|$, $i = 0, \ldots, l-1$, then it is easy to see that the total cost for the computation of the PHSs $\{P^{(i)}\}$ in step (PH.6) along with the coefficients of the residuals $\{R^{(i)}\}$ required by INITIAL_PH is approximately

$$2(k+1)\mu_i \nu_i. \tag{6.2}$$

Theorem 6.1. *The algorithm PADE_HERMITE requires approximately*

$$\mathrm{O}\big((k+1)\| n \|^2 + (k+1)^2 s^2 \| n \|\big) \tag{6.3}$$

multiplications in F, where $s = \max(s_0, s_1, \ldots, s_{l-1})$ *and* $s_i = m_0^{(i+1)} - m_0^{(i)} - 1$ *is the ith stepsize.*

Proof. Equations (6.1) and (6.2) imply that the asymptotic cost of PADE_HERMITE is given by

$$\sum_{i=0}^{l-1} \left[\nu_i^3 + (k+1)\nu_i^2 + \nu_i \mu_i (k+1) \right], \tag{6.4}$$

where $\nu_i \leqslant s_i(k+1) \leqslant s(k+1)$, and

$$\sum_{i=0}^{l-1} \nu_i \leqslant \mu_l \leqslant \| n \|. \tag{6.5}$$

Therefore the cost of the algorithm is bounded approximately by

$$\begin{aligned}
&(k+1)^2 s^2 \sum_{i=0}^{l-1} \nu_i + (k+1)\| n \| \sum_{i=0}^{l-1} \nu_i + (k+1) \sum_{i=0}^{l-1} \nu_i \mu_i \\
&\leqslant (k+1)^2 s^2 \| n \| + (k+1)\| n \|^2 + (k+1) \sum_{i=0}^{l-1} (\mu_{i+1} - \mu_i)\mu_i \\
&\leqslant (k+1)^2 s^2 \| n \| + (k+1)\| n \|^2 + (k+1) \sum_{j=0}^{\| n \|} (j+1-j)j \\
&\leqslant \mathrm{O}\big((k+1)^2 s^2 \| n \| + (k+1)\| n \|^2\big). \qquad \square
\end{aligned} \tag{6.6}$$

Note that the second term in the cost complexity expression (6.3) accounts for the costs arising from all invocations of INITIAL_PH, whereas the first term accounts for all the other costs. Generally speaking, if a large step s_i is required by PADE_HERMITE, then s is large and the second term in (6.3) dominates, whereas, if all step sizes s_i are small, then the first term dominates.

Example 6.2. In the perfect case, $s_i = 1$ for all i. In this case the second term in (6.3) becomes $\mathrm{O}((k+1)^2 \| n \|)$ and so the complexity of the algorithm becomes $\mathrm{O}((k+1)\| n \|^2)$. Similarly, when the vector of power series is *near-perfect* (cf. [19]), then the s_i are bounded by a fixed constant, so again the second term in (6.3) dominates. At the other extreme, when all points with the possible exception of the last along the computational path are singular, that is, $s = s_0 = \max(n_j + 1)$ and $(k+1)s \geqslant \| n \|$, then the second term in (6.3) becomes $\mathrm{O}(\| n \|^3)$ which corresponds to the cost of Gaussian elimination of the full generalized Sylvester system (2.7).

The first term in (6.3) becomes irrelevant here; indeed, the solution returned by the algorithm is exactly that obtained by the first invocation of INITIAL_PH in step (PH.3).

Example 6.3. When $k=1$, it can be shown that the matrices appearing in the INITIAL_PH algorithm are always triangular, hence the cost of steps (PH.3) and (PH.7) are reduced to $2\nu_0^2$ and $2\nu_i^2$, respectively. The corresponding total cost of determining a Padé approximant of type (m, n) in this case is then bounded by $O((m+n)^2)$. This is the case regardless of any assumptions on the size of the steps from one nonsingular node to the next. Gaussian elimination would require $O((m+n)^3)$ operations in this case.

When $n_0 = \cdots = n_k$, Example 6.2 shows that the complexity of PADE_HERMITE when the vector of power series is perfect is $O((k+1)N^2)$, where $N=(k+1)(n_0+1)$ is the size of the associated Sylvester matrix. This agrees with the results of [15] under the same assumptions. In the nonperfect case, however, their algorithm breaks down and so a method such as Gaussian elimination, with a cost of $O(N^3)$ operations, is required. With the use of PADE_HERMITE, even the existence of only one nonsingular point along the diagonal can result in significant speedup.

References

[1] A.C. Antoulas, On recursiveness and related topics in linear systems, *IEEE Trans. Automat. Control* **AC-31** (12) (1986) 1121–1135.

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

[3] G.A. Baker and P. Graves-Morris, *Padé Approximants Part II: Extensions and Applications* (Addison-Wesley, Reading, MA, 1981).

[4] A. Bultheel and M. Van Barel, A matrix Euclidean algorithm and matrix Padé approximation, Technical Report, Dept. Comput. Sci., Kath. Univ. Leuven, 1988.

[5] S. Cabay and D.K. Choi, Algebraic computations of scaled Padé fractions, *SIAM J. Comput.* **15** (1986) 243–270.

[6] J. Coates, On the algebraic approximation of functions, I–III, *Nederl. Akad. Wetensch.* **69** (1966) 421–461.

[7] M.G. de Bruin, Some convergence results in simultaneous rational approximation to the set of hypergeometric functions $\{F_1(1; c_i, z)\}$, in: H. Werner and H.J. Bunger, Eds., *Padé Seminar 1983* (Springer, Berlin, 1983) 95–117.

[8] J. Della Dora and C. Dicrescenzo, Approximants de Padé–Hermite, *Numer. Math.* **43** (1984) 23–57.

[9] W.B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, *SIAM Rev.* **14** (1972) 1–61.

[10] P. Graves-Morris, Vector-valued rational interpolants II, *IMA J. Numer. Anal.* **4** (1984) 209–224.

[11] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators* (Birkhäuser, Basel, 1984).

[12] M.Ch. Hermite, Sur l'expression $U \sin x + V \cos x + W$, extrait d'une lettre a M. Paul Gordan, *Oeuvres Tome III* (1873) 135–145.

[13] M.Ch. Hermite, Sur quelques approximations algébriques, extrait d'une lettre a M. Borchardt, *Oeuvres Tome III* (1873) 146–149.

[14] H. Jager, A multidimensional generalization of the Padé table I–VI, *Nederl. Akad. Wetensch.* **67** (1964) 192–249.

[15] T. Kailath, S.-Y. Kung and M. Morf, Displacement ranks of matrices and linear equations, *J. Math. Anal. Appl.* **68** (1979) 395–407.

[16] G. Labahn and S. Cabay, Matrix Padé fractions and their computation, *SIAM J. Comput.* **18** (4) (1989) 639–657.

[17] G. Labahn, D.K. Choi and S. Cabay, Inverses of block Hankel and block Toeplitz matrices, *SIAM J. Comput.* **19** (1) (1990) 98–123.

[18] K. Mahler, Zur Approximation der Exponentialfunktion und des Logarithmus, Teil I, *J. Reine Angew. Math.* **166** (1932) 118–136.
[19] K. Mahler, Perfect systems, *Compositio Math.* **19** (1968) 95–166.
[20] R.J. McEliece and J.B. Shearer, A property of Euclid's algorithm and an application to Padé approximation, *SIAM J. Appl. Math.* **34** (1978) 611–617.
[21] H. Padé, Sur la représentation approchée d'une fonction par des fractions rationelles, *Ann. Ecole Nor.* (1892) 1–93 (supplement Thesis).
[22] S. Paszkowski, Recurrence relations in Padé–Hermite approximation, *J. Comput. Appl. Math.* **19** (1) (1987) 99–107.
[23] R.E. Shafer, On quadratic approximation, *SIAM J. Numer. Anal.* **11** (1974) 447–460.