The computation of non-perfect Padé-Hermite approximants

Marc Van Barel and Adhemar Bultheel
Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A
B-3001 Leuven (Belgium)

May 22, 2002

Abstract

We describe a simple and efficient algorithm to generate a number of polynomial vectors which can be used to describe all possible solutions for a type I Padé-Hermite problem. If σ denotes the order of approximation, which is a measure for the size of the Padé-Hermite problem, it uses only order σ^2 operations, even if the given system is not perfect. To this end, the problem is considered as a special case of a generalized Padé-Hermite problem which is also defined and analysed.

Keywords : Padé-Hermite approximants, Padé-Hermite forms, type I polynomials, Latin polynomial problem, non perfect system, non normal Padé-Hermite table

AMS(MOS) Classification : Primary : 41A21, Secondary : 41A20, 65D15

The computation of non-perfect Padé-Hermite approximants

Marc Van Barel and Adhemar Bultheel
Department of Computer Science
K.U. Leuven

May 22, 2002

Abstract

We describe a simple and efficient algorithm to generate a number of polynomial vectors which can be used to describe all possible solutions for a type I Padé-Hermite problem. If σ denotes the order of approximation, which is a measure for the size of the Padé-Hermite problem, it uses only order σ^2 operations, even if the given system is not perfect. To this end, the problem is considered as a special case of a generalized Padé-Hermite problem which is also defined and analysed.

1 Introduction

More than a hundred years ago, Hermite introduced, what we call now, Padé-Hermite or Hermite-Padé approximations to a set of functions [23, 21, 22]. He, already, made a distinction between the Latin or type I and the German or type II polynomial problem. The Latin and German refers to the font type that was originally used in the work of Mahler, Jager and Coates to denote the polynomials. The type I polynomials were denoted by a Roman font and the type II polynomials by a Gothic font. Jager [25] actually uses the terms Latin and German polynomials.

Nowadays, the notion of Padé-Hermite approximation is mostly applied for type I polynomial approximations, while the type II problem is in fact studied as a simultaneous Padé or vector-Padé approximation problem. Hermite used these approximations to prove the irrationality of the number e. The theory was further developed by Hermite and Padé [34, 24, 35] as a generalization of the Padé approximation problem [33]. This explains the confusion in calling them Padé-Hermite or Hermite-Padé approximants. Since the problem was studied by people who approached it as a generalization of a Padé approximation problem, they called it Padé-Hermite approximants. However, historically, Hermite's name should come first.

We have not the intention to make a complete overview of the relevant bibliography here. Let us just mention some references. As usual in this kind of Padé-like theory, some papers are in the field of the Padé-Hermite theory as such, while others have strong roots in number theory or in the approximation of functions.

The early references consider especially the use of Padé-Hermite approximants to approximate very specific (systems of) function(s) and numbers and this work is still going on. We mention from the earlier papers of this century the work of Mahler [29, 30, 28], from the 1960's a paper by Coates [12], and from the last twenty years Shafer [40], de Bruin [16, 15, 14], Chudnovsky [8, 9, 11, 10] and Borwein [4, 5]. In this respect, the study of the normality or perfectness of the system of functions is an important issue. In that case, many interesting properties of the polynomials of both types and connections between them can be proved.

Another interesting topic in the evolution of Padé-Hermite theory and vector Padé theory was the introduction of a Padé-Hermite table. It has been studied extensively. We refer to the basic work of Mahler [31] and Jager [25] and more recent work by de Bruin [13], Nikišin [32], Della Dora and di Crescenzo [17, 18, 19, 20] and Paszkowski [36, 37, 38, 39].

Instead of taking the interpolation conditions all at the point $z = 0$, we can distribute these conditions between several points. This generalization was introduced by Mahler [31], Jager [25] and Coates [12]. Lübbe [27] coined the term M-Padé approximant. These approximants were further studied by Beckermann [1, 3, 2].

After this small introduction to the history of the problem, let us explain the purpose of this paper. We don't want to study the singular structure of the Padé-Hermite table because Beckermann [1, 3] describes this structure for the more general M-Padé approximation table. This extends the results of Della Dora [18] and Paszkowski [39]. Instead, we want to develop an elegant and fast algorithm to compute all the solutions of the Padé-Hermite approximation problem following a "diagonal" in the Padé-Hermite table. To clarify the paper, we only investigate the Padé-Hermite approximation problem, where all the interpolation points are confluent and equal to zero. However, also the more general M-Padé approximation problem can be solved in a simple and efficient way using a very similar algorithm to the one given here. This generalization will be developed in a future publication.

Almost all of the work on the algorithmic aspects done previously handles the case where the Padé-Hermite table is normal. We mention the recurrence relations of Paszkowski [38] and those given by Della Dora and di Crescenzo [20] as typical examples. In this normal case, the number of operations is of the order σ^2 where σ is the order of approximation.

In [42], we gave a nice and fast algorithm solving the type II or German polynomial problem for arbitrarily chosen interpolation points even when the simultaneous Padé approximation problem is non-normal (see also [7]). In the present paper, we want to develop a similar algorithm to solve the type I or Latin polynomial problem using order σ^2 operations in all circumstances even for the non-normal or non-perfect Padé-Hermite table. We want to do better than for example in [26], where a square matrix Padé approximation problem is solved with order σ^2 operations in the normal case but in some non-normal cases σ^3 operations are required. As far as we know, only Beckermann [2] has also developed an order σ^2 algorithm working also for a non-normal table. In the sequel, we shall indicate the similarities and the differences with the method developed here.

The paper can be summarized as follows: in section 2, we pose the Padé-Hermite approximation problem and embed it in a more general setting. In section 3, we prove that a basis for all solutions can be obtained using shift chains of so called fundamental vectors. In section 4, a recursive algorithm is developed to compute these fundamental vectors. Finally, in section 5, we give an example of these computations and we illustrate how we can easily implement additional constraints on the Padé-Hermite forms.

2 The Padé-Hermite approximation problem

In this section we define the type I or Latin polynomial problem which we call here Padé-Hermite problem for simplicity. The polynomials and formal series may have coefficients from some arbitrary field $\mathbb{F}$.

Definition 1 (Padé-Hermite form) *Given (n+1) formal power series $f_i, i = 0, 1, \ldots, n$ and an $(n + 1)$-vector $(\mu_0, \mu_1, \ldots, \mu_n)$ of integer numbers $\mu_i \geq -1$ (at least one of them is ≥ 0), a Padé-Hermite form of type $(\mu_0, \mu_1, \ldots, \mu_n)$ is defined as a non-trivial polynomial $(n + 1)$-tuple $(p_0, p_1, \ldots, p_n)$ satisfying*

$$\sum_{j=0}^{n} p_j(z) f_j(z) = O(z^\sigma) \tag{1}$$

with

$$\deg p_j(z) \leq \mu_j, \quad j = 0, 1, \ldots, n \tag{2}$$

and

$$\sigma = \sum_{j=0}^{n} \mu_j + n. \tag{3}$$

The condition of non-triviality excludes the zero $(n+1)$-tuple to be a Padé-Hermite form. The notation in the right hand side of (1) means that this formal series starts of with the term in z^σ. When one of the integer numbers $\mu_j = -1$, condition (2) becomes $\deg p_j(z) \leq \mu_j = -1$, i.e. p_j is the zero polynomial. Hence, the formal power series f_j does not play a role in condition (1). In other words, f_j and p_j can be dropped from the problem setting resulting in a Padé-Hermite form problem with n formal power series instead of $(n + 1)$. At least one of the integer numbers μ_j has to be greater than or equal to zero to guarantee the existence of a non-trivial solution. In general, the interpolation conditions need not be given all at the point $z = 0$. There could be given more general interpolation conditions in several interpolation points. We refer to the work of Mahler [31], Jager [25] and Coates [12], and to the more recent work of Lübbe [27] and Beckermann [1, 3, 2]. We shall exclude this case for simplicity.

The polynomials p_j are also called Latin or type I polynomials. In [42], we solve the dual problem, the German or type II polynomial approximation problem. There we consider the more general case of a vector rational approximation problem where some of the interpolation points may or may not be confluent. In an extreme case, all the interpolation points are confluent and equal to zero. Then we get the simultaneous Padé approximation problem.

To simplify the notation, we shall in the sequel denote $(n+1)$-tuples in bold face letters. They are to be considered as row vectors. Otherwise, they will be explicitly defined as column vectors, e.g. the $(n + 1)$-tuple $\boldsymbol{f}$ of formal power series is a column vector. Operations and inequalities are to be understood componentwise, except the inner product between a row and column vector. For example, the notation $\boldsymbol{\mu} \geq -\mathbf{1}$ indicates that each component of the row vector $\boldsymbol{\mu}$ is greater than or equal to -1, whereas $\boldsymbol{\mu} \neq -\mathbf{1}$ says that at least one component of $\boldsymbol{\mu}$ is different from -1. Note that $\mathbf{1} = (1, 1, \ldots, 1)$. As a further illustration, we reformulate the previous definition as follows.

Definition 2 (Padé-Hermite form in compact notation) *Given an $(n+1)$-tuple $\boldsymbol{f}(z)$ of formal power series, and an $(n + 1)$-tuple $\boldsymbol{\mu}$ of integers $-\mathbf{1} \neq \boldsymbol{\mu} \geq -\mathbf{1}$, we call an $(n + 1)$-tuple of polynomials $\boldsymbol{p}(z)$ a Padé-Hermite form of type $\boldsymbol{\mu}$ if*

$$R(z) = \boldsymbol{p}(z)\boldsymbol{f}(z) = O(z^\sigma), \tag{4}$$

$$\sigma - n = |\boldsymbol{\mu}| \stackrel{\text{def}}{=} \sum_{i=0}^{n} \mu_i, \tag{5}$$

$$\deg \boldsymbol{p}(z) \leq \boldsymbol{\mu}. \tag{6}$$

The notation $R(z) = O(z^\sigma)$ means that $\left(\frac{d}{dz}\right)^k R(z)$ is zero in $z = 0$ for $k = 0, 1, \ldots, \sigma - 1$. The formal power series $R(z)$ will play an important role in the sequel.

Definition 3 (residual series) *Given an $(n+1)$-tuple $\boldsymbol{f}(z)$ of formal power series, the residual series $R(z)$ of an $(n+1)$-tuple of polynomials $\boldsymbol{p}(z)$ with respect to $\boldsymbol{f}(z)$ is defined as*

$$R(z) = \boldsymbol{p}(z)\boldsymbol{f}(z).$$

The Padé-Hermite approximation problem is to find some or all Padé-Hermite forms of a certain specified type. Note that all the solutions of such a problem (extended with the zero vector) form a vector space of dimension at least one because the number of unknowns $\sigma + 1$ is one more than the number of homogeneous linear conditions (1) or, equivalently (4).

To solve the Padé-Hermite approximation problem, we shall embed it in a more general one. Therefore, we introduce the notion of $\boldsymbol{\tau}$-degree.

Definition 4 ($\boldsymbol{\tau}$-degree) *Given an $(n+1)$-vector $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_n)$ of integer numbers, the $\boldsymbol{\tau}$-degree of a polynomial $(n+1)$-tuple is defined as*

$$\boldsymbol{\tau}\text{-deg}\,(p_0, p_1, \ldots, p_n) = \max\{\deg p_j - \tau_j; j = 0, \ldots, n\}.$$

We define the generalized Padé-Hermite problem as follows.

Definition 5 (generalized Padé-Hermite problem) *Given the $(n+1)$-tuple of formal power series $\boldsymbol{f}$, the $(n+1)$-tuple of integer numbers $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_n)$, the integer number δ and the non negative integer $\sigma \geq 0$, the generalized Padé-Hermite problem of type $\boldsymbol{\tau}$ for order σ and degree δ is to find all polynomial $(n+1)$-tuples $\boldsymbol{p} = (p_0, p_1, \ldots, p_n)$ satisfying*

$$\boldsymbol{p}(z)\boldsymbol{f}(z) = O(z^\sigma), \tag{7}$$
$$\boldsymbol{\tau}\text{-deg}\,\boldsymbol{p} \leq \delta. \tag{8}$$

As before, we can note that the set of all solutions of a generalized Padé-Hermite problem for a fixed type vector $\boldsymbol{\tau}$ and given δ and σ forms a vector space of finite dimension. We shall usually call it solution space and denote it by S_δ^σ.

Let us connect the generalized Padé-Hermite problem to the genuine Padé-Hermite problem. We can consider an arbitrary integer k. With $\boldsymbol{\tau} = \boldsymbol{\mu} + \boldsymbol{k} = \boldsymbol{\mu} + (k, k, \ldots, k)$, the generalized Padé-Hermite form is identical to the genuine Padé-Hermite form, that we defined before, if we put $\sigma = |\boldsymbol{\mu}| + n$ and $\delta = -k$. Thus S_{-k}^σ with $\boldsymbol{\tau} = \boldsymbol{\mu} + \boldsymbol{k}$ and $\sigma = |\boldsymbol{\mu}| + n$ is the space of all Padé-Hermite forms of type $\boldsymbol{\mu}$ (plus the zero vector). Note that for the special case $\boldsymbol{\tau} = \boldsymbol{\mu}$, we get equivalence when $\delta = 0$ and $\sigma = |\boldsymbol{\mu}| + n$.

3 A basis for S_δ^σ

For a fixed type $\boldsymbol{\tau}$ and order $\sigma \geq 0$, we can construct a basis for each of the solution spaces $S_\delta^\sigma, \delta = -\infty, \ldots, +\infty$. We shall characterize this basis as the union of shift chains. Such a shift chain will consist of a fundamental vector (which is a certain polynomial $(n+1)$-tuple) and its shifted versions, i.e. successive multiplications with z. There are $n+1$ fundamental vectors, hence there can be at most $n+1$ such shift chains in the basis. If $\boldsymbol{a}_j$ is the j-th fundamental vector ($0 \leq j \leq n$), then a shift chain for $\boldsymbol{a}_j$ has the form $\{\boldsymbol{a}_j(z), z\boldsymbol{a}_j(z), \ldots, z^\gamma \boldsymbol{a}_j(z)\}$. If $\gamma < 0$, we say that the chain is empty. We denote the components of $\boldsymbol{a}_j$ as $\boldsymbol{a}_j = (a_{j,0}, a_{j,1}, \ldots, a_{j,n})$.

The set of $n+1$ fundamental vectors are polynomial solutions of equation (7), with no restriction on the degree δ, but which are independent and minimal in a specific way. The folowing definition doesn't make it obvious that such a set of fundamental vectors always exists, however it is a most practical definition to prove the theorem about the basis of S_δ^σ. That they indeed exist will follow from the construction by the algorithm in the next section.

Definition 6 (fundamental vectors) *Given the* $(n+1)$ *formal power series* $f_j, j = 0, 1, \ldots, n$, *the type* $\boldsymbol{\tau}$ *and the order* $\sigma > 0$. *We call* $\boldsymbol{a}_j, j = 0, 1, \ldots, n$ *a set of fundamental vectors if they satisfy the following conditions:*

1. *They solve equation (7).*
2. *They are linearly independent over the ring of polynomials. This means that there exist no polynomials* $q_j, j = 0, 1, \ldots, n$, *except the zero polynomials, such that*
$$\sum_{j=0}^{n} q_j(z)\boldsymbol{a}_j(z) \equiv \mathbf{0}.$$
3. $\boldsymbol{a}_1(0), \boldsymbol{a}_2(0), \ldots, \boldsymbol{a}_n(0)$ *are linearly independent over the field* $\mathbb{F}$ *and the constant coefficient of the polynomial* $(n+1)$*-tuple* $\boldsymbol{a}_0$ *is the zero vector or* $\boldsymbol{a}_0(0) = \mathbf{0} = (0, 0, \ldots, 0)$.
4. $\boldsymbol{a}_0$ *satisfies (7) with strict order, i.e.*
$$\boldsymbol{a}_0(z)\boldsymbol{f}(z) = O_s(z^\sigma) \stackrel{\text{def}}{=} R_{0,\sigma} z^\sigma + O(z^{\sigma+1}) \quad \textit{with} \quad R_{0,\sigma} \neq 0.$$
5. *The sum of the* $\boldsymbol{\tau}$*-degrees of* $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ *is minimal.*

For the special case $\sigma = 0$, *we define the fundamental vectors as the* $(n+1)$ *canonical unit vectors, i.e. the rows of the unit matrix of size* $n+1$.

Notes: Let us give some comments about the strict order condition for $\boldsymbol{a}_0$ and the minimality condition for the $\boldsymbol{\tau}$-degrees of the fundamental vectors.

- The minimality of the $\boldsymbol{\tau}$-degrees of $\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ implies that $\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are $\boldsymbol{\tau}$-row reduced, by this we mean that the matrix
$$\begin{bmatrix} \boldsymbol{a}_0 \\ \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} z^{-\tau_0} & & & \\ & z^{-\tau_1} & & \\ & & \ddots & \\ & & & z^{-\tau_n} \end{bmatrix},$$
is row reduced. By definition, this matrix is called row reduced if the (vector) coefficients of the highest degree terms, of the $n+1$ rows are linearly independent. This is indeed a consequence of the minimality of the $\boldsymbol{\tau}$-degrees because if this matrix was not $\boldsymbol{\tau}$-row reduced, we could easily generate a new set of fundamental vectors using elementary polynomial combinations such that at least one of the fundamental vectors has a smaller $\boldsymbol{\tau}$-degree while the others do not increase. This would contradict the minimality of the given set. For details on this, one may consult [42].

- Taking into account that $\boldsymbol{a}_0(0) = \mathbf{0}$, the strict order condition for $\boldsymbol{a}_0$ can only be satisfied if there exists at least one f_j having strict order ρ, $f_j(z) = O_s(z^\rho)$, with $\rho + 1 \leq \sigma$. From now on, we shall assume this to be true.

  For the special case $\sigma = 0$, this assumption should be replaced by the condition that there exists at least one f_j for which $f_j(0) \neq 0$. We can then choose the j-th unit vector to be $\boldsymbol{a}_0$ and the strict order condition will also be satisfied. All other conditions we had for $\sigma > 0$ will also be satisfied by a permutation of the fundamental vectors for $\sigma = 0$, except of course the condition $\boldsymbol{a}_0(0) = \mathbf{0}$.

With this assumption about the formal power series, we can give a short proof showing the existence of the fundamental vectors of type $\boldsymbol{\tau}$ and order σ. We need the following definition.

Definition 7 (projection) *The projection P_σ of order σ of a formal power series f is defined as the polynomial of maximum degree $\sigma - 1$ satisfying*

$$P_\sigma f(z) = f(z) + O(z^\sigma),$$

i.e. $P_\sigma f$ is the polynomial part of f having maximum degree $\sigma - 1$.

Theorem 1 (existence of the fundamental vectors) *Given the $(n+1)$ formal power series $f_j, j = 0, 1, \ldots, n$, the type $\boldsymbol{\tau}$ and the order σ. Assume that one of the formal power series f_k satisfies $f_k(0) \neq 0$. Then, the fundamental vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ defined by Definition 6 exist.*

Proof. For $\sigma = 0$, the proof is trivial. For $\sigma > 0$, we consider the following polynomial $(n+1)$-tuples:

$$\begin{bmatrix} \boldsymbol{a}_0(z) \\ \boldsymbol{a}_1(z) \\ \boldsymbol{a}_2(z) \\ \vdots \\ \boldsymbol{a}_k(z) \\ \boldsymbol{a}_{k+1}(z) \\ \vdots \\ \boldsymbol{a}_n(z) \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & z^\sigma & 0 & \cdots & 0 \\ P_\sigma f_k & 0 & \cdots & 0 & -P_\sigma f_0 & 0 & \cdots & 0 \\ 0 & P_\sigma f_k & \cdots & 0 & -P_\sigma f_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & P_\sigma f_k & -P_\sigma f_{k-1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -P_\sigma f_{k+1} & P_\sigma f_k & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & -P_\sigma f_n & 0 & \cdots & P_\sigma f_k \end{bmatrix}.$$

This set of polynomial $(n+1)$-tuples satisfies all conditions of Definition 6 except the last one about the minimality of the sum of the $\boldsymbol{\tau}$-degrees. We leave the details as an exercise for the reader. Since we have given here one specific solution satisfying the first four conditions, the set of all solutions is not empty. Therefore, there must be at least one solution, i.e. a set of polynomial $(n+1)$-tuples, with minimal sum of the $\boldsymbol{\tau}$-degrees. Hence, there exists a set of fundamental vectors. □

In the next section we shall give an algorithm which actually constructs the fundamental vectors for some $\sigma + 1$, given the fundamental vectors for σ. Since those for $\sigma = 0$ exist by definition, the existence for every σ follows by induction under the given assumption on $\boldsymbol{f}$.

Before we shall give this algorithm to generate the fundamental vectors, we show their relevance in solving the generalized Padé-Hermite problem. The fundamental vectors can be used to construct a basis for each of the solution spaces S_δ^σ as indicated in Theorem 3 below. In the proof of this theorem, we shall use some of the following relations that exist between the solution spaces. These relations shall also form the basis of our algorithm.

Lemma 2 *For fixed type $\boldsymbol{\tau}$, it holds that*

$$\begin{array}{rcl} S_\delta^\sigma & \subset & S_{\delta+1}^\sigma \\ zS_\delta^\sigma & \subset & S_{\delta+1}^{\sigma+1} \\ S_\delta^{\sigma+1} & \subset & S_\delta^\sigma \\ z\boldsymbol{a} \in S_{\delta+1}^{\sigma+1} & \Rightarrow & \boldsymbol{a} \in S_\delta^\sigma. \end{array}$$

Proof. The proof follows immediately from the definition of the solution spaces into play, or from the nested structure of the homogeneous linear equations implied by (1). □

We are now ready to describe and prove our basis for the solution spaces.

Theorem 3 (a basis for S_δ^σ) *Given the order $\sigma \geq 0$, the degree δ, the type $\boldsymbol{\tau}$ and a set of fundamental vectors $\boldsymbol{a}_i$ having $\boldsymbol{\tau}$-degree δ_i. A basis for the solution space S_δ^σ is given by the union of at most $n+1$ shift chains*

$$\bigcup_{i=0}^{n} \{\boldsymbol{a}_i(z), z\boldsymbol{a}_i(z), \ldots, z^{\delta-\delta_i}\boldsymbol{a}_i(z)\}.$$

Note that the fundamental vector $\boldsymbol{a}_i$ does not appear in this basis when $\delta_i > \delta$.

Proof. It is easy to see that the theorem is true for $\sigma = 0$. We leave this to the reader.

Suppose that $\sigma > 0$. Consider an element $\boldsymbol{a}$ of the solution space S_δ^σ. We shall show that there is a unique way to write this element as a linear combination of the vectors given above. Which shows that these vectors form a basis.

We first prove that $\boldsymbol{a}(0)$ is always a linear combination of $\boldsymbol{a}_i(0), i = 1, 2, \ldots, n$. Suppose this is not true. Then

$$\boldsymbol{A}(z) = \begin{bmatrix} \boldsymbol{a}(z) \\ \boldsymbol{a}_1(z) \\ \boldsymbol{a}_2(z) \\ \vdots \\ \boldsymbol{a}_n(z) \end{bmatrix} = O(z^0)$$

and $\boldsymbol{A}(0)$ has full rank. Because $\boldsymbol{a}$ and all fundamental vectors satisfy relation (7), we can also write

$$\boldsymbol{A}(z) \begin{bmatrix} f_0(z) \\ f_1(z) \\ \vdots \\ f_n(z) \end{bmatrix} = \boldsymbol{A}\boldsymbol{f} = O(z^\sigma),$$

where the order-relation holds for each component of the vector of the residual series. Bringing $\boldsymbol{A}$ to the other side of this equation, leads to the componentwise order-relation

$$\boldsymbol{f} = O(z^0)O(z^\sigma) = O(z^\sigma).$$

This is in contradiction with the assumption that there exists an f_j satisfying $f_j = O_s(z^\rho)$ with $\rho < \sigma$. Therefore, $\boldsymbol{a}(0)$ is a linear combination of $\boldsymbol{a}_i(0)$. Moreover in this linear combination only those $\boldsymbol{a}_i$ can appear for which $\delta_i \leq \delta$ because the $\boldsymbol{\tau}$-degrees δ_i are minimal by definition of the fundamental vectors. Hence, we can decompose $\boldsymbol{a}$ as $\boldsymbol{a} = \sum_{j=1}^n c_j\boldsymbol{a}_j + \boldsymbol{a}'$ with $\boldsymbol{a}'$ divisible by z and this decomposition is unique. Using the definition of $\boldsymbol{a}_0$, we can write $\boldsymbol{a}'$ in exactly one way as $\boldsymbol{a}' = c_0\boldsymbol{a}_0 z^{\delta-\delta_0} + \boldsymbol{a}''$ with $\boldsymbol{a}''$ divisible by z and in $S_\delta^{\sigma+1}$

$$\boldsymbol{a}''(z)\boldsymbol{f}(z) = O(z^{\sigma+1}).$$

Therefore, $\boldsymbol{a}''(z)/z$ is also a polynomial $(n+1)$-tuple and an element of $S^{\sigma}_{\delta-1}$. The theorem now follows by induction. □

4 A fast recursive algorithm to solve the generalized Padé-Hermite approximation problem

In this section, we shall develop an elegant and fast algorithm to compute a set of fundamental vectors. As a matter of fact, because we can start from the trivial choice of fundamental vectors when $\sigma = 0$, it is sufficient to describe how we can compute the fundamental vectors $\boldsymbol{a}_i^{\sigma+1}$ of order $\sigma+1$ from the fundamental vectors $\boldsymbol{a}_i^{\sigma}$ of the previous order. As we have explained in the previous section, we can take the unit vectors as fundamental vectors for $\sigma = 0$ if we assume that there exists an f_j with $f_j(z) = O_s(z^0)$. This is the only assumption to be made, because we shall actually construct all the fundamental vectors for subsequent orders $\sigma = 1, 2, \ldots$ and thereby also prove their existence, without any further assumption. Throughout this section, the type $\boldsymbol{\tau}$ is fixed. We denote the $\boldsymbol{\tau}$-degree of the fundamental vector $\boldsymbol{a}_i^{\sigma}$ as δ_i^{σ} and the corresponding residual series as R_i^{σ}.

To describe the algorithm, we need the concept of residual which is the leading coefficient in the residual series which may or may not be zero. More precisely, it is defined as folows.

Definition 8 (residual) *The residual $R_{i,\sigma}$ of the fundamental vector $\boldsymbol{a}_i^{\sigma}$ is defined by the relation*

$$R_i^{\sigma}(z) = \boldsymbol{a}_i^{\sigma}\boldsymbol{f} = R_{i,\sigma}z^{\sigma} + O(z^{\sigma}), \tag{9}$$

i.e. the residual is the coefficient of order σ in the corresponding residual series.

From Definition 6, it follows that $R_{0,\sigma} \neq 0$ for $\sigma > 0$, and $R_{j,0} \neq 0$ when $f_j(0) \neq 0$.

First of all, we shall give an informal description of the algorithm. The algorithm starts by taking the fundamental vectors $\{\boldsymbol{a}_i^0; i = 0, \ldots, n\}$ of order 0 equal to the $(n+1)$-unit vectors having $\boldsymbol{\tau}$-degrees $\delta_i^0 = -\tau_i$ and residual series f_i.

Now we describe the recursion step to go from σ to $\sigma+1$. Therefore suppose that we have computed a set of fundamental vectors $\{a_i^{\sigma}; i = 0, \ldots, n\}$ of order σ and want to compute a set of fundamental vectors $\{a_i^{\sigma+1}; i = 0, \ldots, n\}$ of order $\sigma+1$. Because each fundamental vector $\boldsymbol{a}_i^{\sigma+1}$ is by definition an element of $S^{\sigma+1}_{\delta_i^{\sigma+1}}$, Lemma 2 says that $\boldsymbol{a}_i^{\sigma+1}$ is also an element of $S^{\sigma}_{\delta_i^{\sigma+1}}$. Therefore, each fundamental vector $\boldsymbol{a}_i^{\sigma+1}$ can be written as a linear polynomial combination of the fundamental vectors of order σ. Hence also the corresponding residual series $R_i^{\sigma+1}$ can be computed as the same linear polynomial combination of the residual series of order σ.

By the second condition of Definition 6, the fundamental vectors $\boldsymbol{a}_i^{\sigma}$ are independent over the polynomials. Also the fundamental vectors $\boldsymbol{a}_i^{\sigma+1}$ should satisfy this condition. This implies that there must be a polynomial matrix $\boldsymbol{V}^{\sigma+1}$ whose determinant may not vanish identically and such that

$$\boldsymbol{A}^{\sigma+1} = \begin{bmatrix} \boldsymbol{a}_0^{\sigma+1} \\ \boldsymbol{a}_1^{\sigma+1} \\ \vdots \\ \boldsymbol{a}_n^{\sigma+1} \end{bmatrix} = \boldsymbol{V}^{\sigma+1}\boldsymbol{A}^{\sigma} = \boldsymbol{V}^{\sigma+1} \begin{bmatrix} \boldsymbol{a}_0^{\sigma} \\ \boldsymbol{a}_1^{\sigma} \\ \vdots \\ \boldsymbol{a}_n^{\sigma} \end{bmatrix}.$$

The following description of this transition matrix $\boldsymbol{V}^{\sigma+1}$ defines the algorithm.

The fundamental vector $\boldsymbol{a}_0^{\sigma+1}$ has to be in a solution space $S_{\delta+1}^{\sigma+1}$ for some δ and it has to be divisible by z. Following Lemma 2, it should be of the form $\boldsymbol{a}_0^{\sigma+1} = z\boldsymbol{a}$ with $\boldsymbol{a} \in S_\delta^\sigma$. Therefore, $\boldsymbol{a}$ can be written as a linear polynomial combination of the fundamental vectors of order σ. To guarantee that δ is minimal and that the residual $R_{0,\sigma+1} \neq 0$ as required, we can take $\boldsymbol{a}$ equal to $\boldsymbol{a}_r^\sigma$, the fundamental vector of order σ with least $\boldsymbol{\tau}$-degree and with residual $R_{r,\sigma} \neq 0$, i.e. satisfying equation (7) with right-hand side $O_s(z^\sigma)$. This $\boldsymbol{a}_r^\sigma$ will always exist because $\boldsymbol{a}_0^\sigma$ for $\sigma > 0$, and $\boldsymbol{a}_j^\sigma$ for $\sigma = 0$ with $f_j(0) \neq 0$, are among the $\boldsymbol{a}_i^\sigma$ with $R_{i,\sigma} \neq 0$, but it need not be uniquely defined. Hence, a possible choice for the fundamental vector $\boldsymbol{a}_0^{\sigma+1}$ is $z\boldsymbol{a}_r^\sigma$ with residual $R_{0,\sigma+1} = R_{r,\sigma}$ and minimal $\boldsymbol{\tau}$-degree $\delta_0^{\sigma+1} = \delta_r^\sigma + 1$.

Each of the other fundamental vectors of order σ can be used as a component of a fundamental vector of order $\sigma + 1$ with the same $\boldsymbol{\tau}$-degree. More precisely, the remaining fundamental vectors $\boldsymbol{a}_j^{\sigma+1}, j = 1, 2, \ldots, n$, of order $\sigma + 1$ are equal to $\boldsymbol{a}_i^\sigma - \boldsymbol{a}_r^\sigma R_{i,\sigma}/R_{r,\sigma}$, $i = 0, 1, \ldots, n$ and $i \neq r$, having $\boldsymbol{\tau}$-degrees $\delta_j^{\sigma+1} = \delta_i^\sigma$. From this recursive construction, it follows immediately that the vectors $\boldsymbol{a}_i^{\sigma+1}$, $i = 0, 1, \ldots, n$ are fundamental vectors of order $\sigma + 1$.

We summarize this method as algorithm *PadeHermite given below. Besides the notation of the square polynomial matrix $\boldsymbol{A}^\sigma$ that we introduced above, we also gather the residual series and the $\boldsymbol{\tau}$-degrees.*

$$\mathcal{R}^{\times\sigma} = \begin{bmatrix} R_0^\sigma \\ R_1^\sigma \\ \vdots \\ R_n^\sigma \end{bmatrix} \quad \textit{and} \quad \boldsymbol{\delta}^\sigma = \begin{bmatrix} \delta_0^\sigma \\ \delta_1^\sigma \\ \vdots \\ \delta_n^\sigma \end{bmatrix}.$$

The polynomial matrix $\boldsymbol{V}^{\sigma+1}$ summarizes the operations to go from order σ to order $\sigma + 1$. The r-th unit vector is denoted as $\boldsymbol{e}_r = [0 \ldots 0\ 1\ 0 \ldots 0]^T$ where the one is the $(r+1)$-st component and superscript T denotes transpose.

Algorithm 1: *PadeHermite*

{ Initialization }

$$\left[\ \boldsymbol{A}^0 \ \mid\ \mathcal{R}^{\times 0}\ \right] = \left[\ \boldsymbol{I}\ \mid\ \boldsymbol{f}\ \right] \quad \text{with} \quad \boldsymbol{\delta}^0 = \begin{bmatrix} -\tau_0 \\ -\tau_1 \\ \vdots \\ -\tau_n \end{bmatrix}.$$

for $\sigma = 0, 1, 2, \ldots$

Determine r as $\delta_r^\sigma = \min\{\delta_j^\sigma | R_{j,\sigma} \neq 0\}$.

Define the matrix $\boldsymbol{V}^{\sigma+1}$

$$\boldsymbol{V}^{\sigma+1} = \begin{bmatrix} 0 & \ldots & 0 & z & 0 & \ldots & 0 \\ 1 & \ldots & 0 & -R_{0,\sigma}/R_{r,\sigma} & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & -R_{r-1,\sigma}/R_{r,\sigma} & 0 & \ldots & 0 \\ 0 & \ldots & 0 & -R_{r+1,\sigma}/R_{r,\sigma} & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & -R_{n,\sigma}/R_{r,\sigma} & 0 & \ldots & 1 \end{bmatrix}.$$

Go from order σ to order $\sigma + 1$ by

$$\left[\ \boldsymbol{A}^{\sigma+1}\ \mid\ \mathcal{R}^{\times \sigma+1}\ \right] = \boldsymbol{V}^{\sigma+1}\left[\ \boldsymbol{A}^{\sigma}\ \mid\ \mathcal{R}^{\times \sigma}\ \right] \quad \text{and} \quad \boldsymbol{\delta}^{\sigma+1} = \P(\boldsymbol{\delta}^\sigma + \boldsymbol{e}_r)$$

with the permutation matrix $\P$ equal to

$$\P = [\boldsymbol{e}_r\ \boldsymbol{e}_0\ \boldsymbol{e}_1 \cdots \boldsymbol{e}_{r-1}\ \boldsymbol{e}_{r+1}\ \boldsymbol{e}_{r+2} \cdots \boldsymbol{e}_n]^T.$$

endfor

Notes:

- *The residuals needed in the algorithm can immediately be read off from the residual series* $\mathcal{R}^{\times\sigma}$ *as the coefficients of order* σ*. We could also drop the updating of the residual series from the algorithm and compute the residuals using the Definition 8, i.e. using an inner product formula (9). This however is more difficult to parallelize. For more information, we refer the interested reader to the work of Bruckstein and Kailath [6] on layer peeling and layer adjoining variants of similar algorithms. For the vector rational interpolation problem, a similar discussion can be found in [42].*

- *The algorithm PadeHermite can be implemented in several ways. When we look at the operations involved, summarized in the matrix* $\boldsymbol{V}^{\sigma+1}$ *we conclude that to compute the components of the* $\boldsymbol{a}_i^{\sigma+1}$*, the only elementary operations required are a shift and an addition of a constant multiple of a polynomial to another one. Hence, instead of using the classical basis* $1, z, z^2, \ldots$ *to represent these component polynomials, we can use any basis as long as these two elementary operations can be implemented efficiently. One possibility is to use orthogonal polynomials* $\phi_0(z), \phi_1(z), \ldots$ *which satisfy a three term recurrence relation* $\phi_k(z) = \lambda_k(z-\alpha_k)\phi_{k-1}(z) - \beta_k\phi_{k-2}(z)$*. From the latter, we see that the shift operation can be efficiently implemented. In our previous publications ([7, 41] and others), we have always developed algorithms using only these two basic operations on polynomials. Therefore, the same reasoning can be made for the representation of the polynomial vectors emerging in all these algorithms.*

- *When we apply this elegant algorithm to construct the Padé-Hermite form of type* τ*, we follow a "diagonal" path in the Padé-Hermite table (which could be non-normal or non-perfect). Indeed, at level* $\sigma = |\boldsymbol{\tau}| + n$ *we compute the fundamental vectors from which all solutions of the Padé-Hermite approximation problem of type* $\boldsymbol{\tau}$ *can be constructed as the elements of* S_0^σ*. At level* $\sigma+n+1$*, all solutions of the Padé-Hermite approximation problem of type* $\boldsymbol{\tau}+\mathbf{1}$ *are the elements of* $S_1^{\sigma+n+1}$*. In general,* $S_k^{\sigma+k(n+1)}$ *is the set of all solutions of the Padé-Hermite problem of type* $\boldsymbol{\tau}+k\mathbf{1}$*. (Pay attention when going backwards: one of the components of* $\boldsymbol{\tau}-k\mathbf{1}$ *can become less than zero.)*

- *This recursive algorithm is fast. Suppose we want to compute the fundamental vectors of order* $\sigma = 0, 1, 2, \ldots, \sigma_{max}$*, i.e. the* **for***-loop in algorithm PadeHermite runs from* $\sigma = 0$ *until* $\sigma = \sigma_{max} - 1$*. In this case we do not need to compute* all *the coefficients of the residual series but only the coefficients which are needed in the subsequent steps of the algorithm. More precisely, when the* **for***-loop index is equal to* σ*, we know that all residual series* $R_i^{\sigma+1}$*,* $i = 0, 1, \ldots, n$ *will satisfy*

$$R_i^{\sigma+1} = O(z^{\sigma+1}).$$

Therefore, we only have to compute the coefficients connected to the powers $\sigma+1, \sigma+2, \ldots, \sigma_{max}-1$ *of* z*. For the same* **for***-loop index* σ*, each of the polynomial components of the square* $(n+1)\times(n+1)$ *matrix* $\boldsymbol{A}^{\sigma+1}$ *has a degree smaller than or equal to* σ *except the components of the first row of* $\boldsymbol{A}^{\sigma+1}$ *which could be of degree* $\sigma+1$*. Let us investigate the operations to go from order* σ *to order* $\sigma+1$*. We shall only indicate the arithmetic operations (in the arbitrary field* $\mathbb{F}$*). To compute the first fundamental vector* $\boldsymbol{a}_0^{\sigma+1}$ *with corresponding residual series* $R_0^{\sigma+1}$*, we only apply a shift operation involving at most* $(n+1)(\sigma+2)$ *coefficients of all the components of* $\boldsymbol{a}_0^{\sigma+1}$ *and* $(\sigma_{max}-1)-\sigma$ *coefficients of* $R_0^{\sigma+1}$*. For each of the remaining fundamental vectors* $\boldsymbol{a}_i^{\sigma+1}$ *with corresponding residual series* $R_i^{\sigma+1}$*,* $i = 1, 2, \ldots, n$*, we have to compute first* $r_i = R_{i-1,\sigma}/R_{r,\sigma}$ *for* $i \leq r$

and $r_i = R_{i,\sigma}/R_{r,\sigma}$ *for* $i > r$*, requiring one division operation. Secondly, we apply one multiplication and one subtraction operation for at most* $(n+1)(\sigma+1)$ *coefficients of all components of* $\boldsymbol{a}_i^{\sigma+1}$ *and* $(\sigma_{max}-1)-\sigma$ *coefficients of* $R_i^{\sigma+1}$*. We neglect the operations to determine* r *and the updating operations on the* $\boldsymbol{\tau}$*-degree vector* $\boldsymbol{\delta}^{\sigma+1}$*. Hence, to go from order* σ *to order* $\sigma+1$*, we need* n *division operations,* $n((n+1)(\sigma+1)+(\sigma_{max}-1)-\sigma)$ *multiplication and addition operations. Therefore, algorithm PadeHermite computes the fundamental vectors of order* $\sigma = 0, 1, 2, \ldots, \sigma_{max}$ *using* $n\sigma_{max}$ *division operations and at most*

$$\sum_{\sigma=0}^{\sigma_{max}-1} n((n+1)(\sigma+1) + (\sigma_{max}-1) - \sigma) = (n^2/2 + n)\sigma_{max}^2$$

multiplication and addition operations, i.e. using a number of arithmetic operations of order σ_{max}^2*, even when the Padé-Hermite table is non-perfect.*

When we know the factors r_i*, all coefficients of all components of the fundamental vectors* $\boldsymbol{a}_i^{\sigma+1}$*,* $i = 1, 2, \ldots, n$*, can be computed simultaneously. Therefore, if we have order* σ_{max} *processors, the algorithm can be executed in parallel in order* σ_{max} *time. Let us give an example. Suppose we have* σ_{max} *processors with* $\sigma_{max} \geq n$*. Using* n *of these processors, we can first compute the* n *factors* r_i *in the time required to do one division operation. To go from order* σ *to order* $\sigma + 1$*, we can secondly use* $\sigma+1 \leq \sigma_{max}$ *processors to compute all coefficients of all components of the fundamental vectors* $\boldsymbol{a}_i^{\sigma+1}$*,* $i = 1, 2, \ldots, n$*, in the time required to execute* $n(n+1)$ *multiplications and* $n(n+1)$ *additions. Using* $(\sigma_{max}-1)-\sigma < \sigma_{max}$ *processors in parallel, we can compute all coefficients of all residual series* $R_i^{\sigma+1}$*,* $i = 1, 2, \ldots, n$*, in the time required to do* n *multiplications and* n *additions. When we neglect the time to determine* r*, to derive* $\boldsymbol{a}_0^{\sigma+1}$ *and* $R_0^{\sigma+1}$ *using shift operations and to update the* $\boldsymbol{\tau}$*-degree vector* $\boldsymbol{\delta}^{\sigma+1}$*, we get a total execution time of* σ_{max} *divisions,* $(n^2+2n)\sigma_{max}$ *multiplications and* $(n^2+2n)\sigma_{max}$ *additions. This execution time can be decreased further. For example, if we can use* $n((n+1)(\sigma+1)+(\sigma_{max}-1)-\sigma) \leq (n^2+n)\sigma_{max}$ *processors in parallel, we get a total execution time of* σ_{max} *divisions,* σ_{max} *multiplications and* σ_{max} *additions.*

Instead of using parallel processors, we could easily run this algorithm on a vector processor. If we store all coefficients of all components of the fundamental vector $\boldsymbol{a}_i^{\sigma+1}$ *and all coefficients of the residual series* $R_i^{\sigma+1}$ *in one big vector, we get* $(n+1)$ *of such coefficient vectors. The arithmetic operations involved can be interpreted as operations with the coefficient vectors. To go from order* σ *to order* $\sigma+1$*, a shift operation on a coefficient vector and* n *additions of a coefficient vector multiplied by a scalar to another coefficient vector are needed, together with* n *scalar divisions to compute the scalar factors* r_i*. Of course, we could use* n *of these vector processors in parallel to speed up the computations.*

- *In [2], Beckermann has developed a recursive algorithm to compute all M-Padé approximants on an "arbitrary staircase" in the M-Padé table, even if this table is non-perfect. The algorithm consists of three steps to go from order* σ *to order* $\sigma+1$*. However, when all interpolation points coincide and when a "diagonal" path is taken in the Padé-Hermite table, the third step of the algorithm can be skipped resulting in an algorithm which is very similar to our algorithm PadeHermite. However, there are some differences. Algorithm PadeHermite computes fundamental vectors satisfying the conditions of Definition 6. These are conditions on the "constant" coefficients, the residual series and the* $\boldsymbol{\tau}$*-degrees of these vectors. Note that the ordering of these fundamental vectors* *is in fact irrelevant. Therefore, the permutation operation connected to the permutation matrix ¶ in algorithm PadeHermite is not necessary. In contrast, the fundamental vectors defined by Beckermann have to satisfy conditions on the "defects" and on the "highest degree" coefficients where the ordering plays an important role. Hence, our fundamental vectors and those described by Beckermann are equivalent but characterized in a different way. The characterization used by Beckermann implies the necessity of a reordering in step 2 of his algorithm and the computation of an integer value κ besides d (called r in our context).*

5 Example

Let us take example 4.1 and 4.4 from Paszkowski [39]. We are given the $n+1=3$ formal power series

$$\begin{aligned} f_0(z) &= -1+z^2+z^5-z^6+z^7+z^8+z^9+\dots \\ f_1(z) &= 1+z-z^5+z^7+2z^8-z^9+\dots \\ f_2(z) &= -z-z^2+2z^6-z^7+z^8+2z^9+\dots \end{aligned}$$

We are looking for all Padé-Hermite forms of type $(1,1,1)$ and type $(2,2,2)$. Hence, the solutions lie on the "diagonal" of the Padé-Hermite table. Therefore, we can choose $\boldsymbol{\tau}=(0,0,0)$, determining the parameter $\boldsymbol{\tau}$ of the $\boldsymbol{\tau}$-degree. In fact, we could have chosen any $\boldsymbol{\tau}$ of the form (k,k,k). Using the fundamental vectors of order $|(k,k,k)|+n$, we can write down all Padé-Hermite forms of type (k,k,k), i.e. in this example the fundamental vectors of order $2,5,8$ are the ones to be used. We use the following notation to show the results of the algorithm at each step

$$\left[\ \boldsymbol{V}^{\sigma}\ \ ||\ \ \boldsymbol{\delta}^{\sigma}\ \ ||\ \ \boldsymbol{A}^{\sigma}\ \ ||\ \ \mathcal{R}^{\times\sigma}\ \right],$$

where we represent $\boldsymbol{V}^{\sigma}$ and $\boldsymbol{A}^{\sigma}$ using the polynomial representation with respect to the classical basis $1,z,z^2,\dots$. The residual series $\mathcal{R}^{\times\sigma}$ are represented coefficient-wise to show the importance of the residuals and the increasing of the order in each step. We summarize the results of the algorithm in table 1.

Notes:

1. *The results of the algorithm allow us to write down a basis of any solution space S_δ^σ of the generalized Padé-Hermite problem of type $\boldsymbol{\tau}=(0,0,0)$ and order $\sigma=0,1,2,\dots,8$ and any degree δ. This basis is constructed using the fundamental vectors and their shifted versions.*

2. *All Padé-Hermite forms of type $(0,0,0)$ are given by the solution space S_0^2 excluding the zero vector. This space is spanned by shift chains of fundamental vectors of order 2 and of ($\boldsymbol{\tau}$-)degree less than 1. There is only one such chain and it contains only one element: the fundamental vector $(1,1,1)$. Thus all Padé-Hermite forms $S_0^2\setminus\{\mathbf{0}\}$ of type $(0,0,0)$ are of the form $c(1,1,1)$ with c a nonzero constant to exclude the trivial solution.*

   The solution space S_1^5 has a basis consisting of two shift chains. The fundamental vector $(1,1,1)$ has $\boldsymbol{\tau}$-degree 0 and its chain contains therefore two elements, while the vector $(1,1-z,0)$ has $\boldsymbol{\tau}$-degree 1 and its chain consists only of the vector itself. Hence we can describe all Padé-Hermite forms of type $(1,1,1)$ explicitly as $S_1^5=\{d(z)(1,1,1)+e(1,-z+1,0)$ with $d(z)=d_0+d_1z\}$ excluding the zero vector.

Table 1: Summary of the results of algorithm *PadeHermite*

$\boldsymbol{V}^\sigma$			$\boldsymbol{\delta}^\sigma$	$\boldsymbol{A}^\sigma$			1	z	z^2	z^3	z^4	z^5	z^6	z^7	z^8	z^9
The results for order 0																
			0	1	0	0	-1	0	1	0	0	1	-1	1	1	1
			0	0	1	0	1	1	0	0	0	-1	0	1	2	-1
			0	0	0	1	0	-1	-1	0	0	0	2	-1	1	2
The results for order 1																
z	0	0	1	z	0	0	0	-1	0	1	0	0	1	-1	1	1
1	1	0	0	1	1	0	0	1	1	0	0	0	-1	2	3	0
0	0	1	0	0	0	1	0	-1	-1	0	0	0	2	-1	1	2
The results for order 2																
0	z	0	1	z	z	0	0	0	1	1	0	0	0	-1	2	3
1	1	0	1	$z+1$	1	0	0	0	1	1	0	0	0	1	4	1
0	1	1	0	1	1	1	0	0	0	0	0	0	1	1	4	2
The results for order 3																
z	0	0	2	z^2	z^2	0	0	0	0	1	1	0	0	0	-1	2
-1	1	0	1	1	$-z+1$	0	0	0	0	0	0	0	0	2	2	-2
0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	4	2
The results for order 4																
z	0	0	3	z^3	z^3	0	0	0	0	0	1	1	0	0	0	-1
0	1	0	1	1	$-z+1$	0	0	0	0	0	0	0	0	2	2	-2
0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	4	2
The results for order 5																
z	0	0	4	z^4	z^4	0	0	0	0	0	0	1	1	0	0	0
0	1	0	1	1	$-z+1$	0	0	0	0	0	0	0	0	2	2	-2
0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	4	2
The results for order 6																
z	0	0	5	z^5	z^5	0	0	0	0	0	0	0	1	1	0	0
0	1	0	1	1	$-z+1$	0	0	0	0	0	0	0	0	2	2	-2
0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	4	2
The results for order 7																
0	0	z	1	z	z	z	0	0	0	0	0	0	0	1	1	4
1	0	-1	5	z^5-1	z^5-1	-1	0	0	0	0	0	0	0	0	-4	-2
0	1	0	1	1	$-z+1$	0	0	0	0	0	0	0	0	2	2	-2
The results for order 8																
z	0	0	2	z^2	z^2	z^2	0	0	0	0	0	0	0	0	1	1
0	1	0	5	z^5-1	z^5-1	-1	0	0	0	0	0	0	0	0	-4	-2
-2	0	1	1	$-2z+1$	$-3z+1$	$-2z$	0	0	0	0	0	0	0	0	0	-10

Similarly, the set of all Padé-Hermite forms of type $(2,2,2)$ *is given as* $S_2^8 = \{g(z)(1-2z, 1-3z, -2z) + h(z^2, z^2, z^2)$ *with* $g(z) = g_0 + g_1 z\}$ *excluding the zero vector. Note that these results are equivalent to those obtained by Paszkowski [39]. However, his results don't show the shift structure of the basis vectors.*

3. *Because the fundamental vectors allow us to describe all solutions of (7-8), we can characterize (generalized) Padé-Hermite forms that satisfy some extra conditions.*

   For example, Paszkowski [39] defines a Padé-Hermite form to be a solution of (1)-(3) with the extra condition that $\boldsymbol{p}(0) \neq 0$. *Once we have computed the fundamental vectors* $\boldsymbol{A}^\sigma$, *and found all solutions described as the linear polynomial combination* $\mathbf{q}(z)\boldsymbol{A}^\sigma(z)$, *then it is easy to pick out the constant coefficient* $\mathbf{q}(0)\boldsymbol{A}^\sigma(0)$ *and put additional conditions upon* $\mathbf{q}(0)$. *Note that, even when we do not compute the fundamental vectors explicitly, we can compute the* $\boldsymbol{A}^\sigma(0)$ *very easily by multiplying the matrices* $\boldsymbol{V}^\sigma(0)$.

   For our example, we get for type $(1,1,1)$ *the additional condition* $d_0(1,1,1)+e(1,1,0) \neq (0,0,0)$ *or equivalently* $d_0 \neq 0$ *and* $d_0 + e \neq 0$. *For type* $(2,2,2)$, *we get the additional condition* $g_0(1,1,0) + h(0,0,0) \neq (0,0,0)$ *or* $g_0 \neq 0$. *Hence, we obtain the same results as Paszkowski in a very simple and efficient way.*

   Another application of this idea gives another result of Paszkowski. He defines the optimal Hermite-Padé form as the Hermite-Padé form satisfying the order relation (7) as far as possible. He proves that this optimal form is unique up to a constant factor. Using our framework of fundamental vectors and corresponding residual series. We can give a constructive proof of this result. We shall not do this explicitly. Instead we shall illustrate the technique for our example. Consider the type $(2,2,2)$ *corresponding to the fundamental vectors of order 8. Looking at the residual series and the* $\boldsymbol{\tau}$*-degrees of the fundamental vectors into play, we see immediately that the maximal order is obtained for the solution* $d_1 z(-2z+1, -3z+1, -2z)$. *Its residual series has maximal order 10. Paszkowski would not consider this as the optimal Padé-Hermite form because it is divisible by* z. *He would say the optimal form does not exist.*

   Another way to get the optimal form is the following. Consider the optimal form of type $(0,0,0)$. *The rule to find the optimal solution is as follows. Since all Padé-Hermite forms of type* $(0,0,0)$ *need fundamental vectors of order 2, we start in the table from the results for order 2. We then search the table for increasing orders* σ *until we find the last order for which there is still a fundamental vector of* $\boldsymbol{\tau}$*-degree* $= 0$. *This is order 6 in our example. Hence, there will be only one fundamental vector of order 6 and of* $\boldsymbol{\tau}$*-degree 0. In the example this is* $(1,1,1)$. *It is the optimal form (up to a constant factor).*

   With respect to the Padé-Hermite form problem, we have drawn our attention especially to the fundamental vectors of orders $|(k,k,k)| + n$. *However, also the fundamental vectors of orders in between can be used to find a minimal Padé-Hermite form, which is yet another criterion for optimality. We could define it as a Padé-Hermite form which has a certain given order* σ *for which the* $\boldsymbol{\tau}$*-degree is minimal. This problem is solved by taking all linear combinations (not polynomial) of the fundamental vectors of order* σ *having minimal* $\boldsymbol{\tau}$*-degree. It is also easy to add the additional condition that the solution should have a non-zero constant coefficient and/or that it should be optimal.*

6 Conclusion

In this paper, we have shown that we can embed the classical Padé-Hermite problem into a generalized Padé-Hermite problem. The solution of this generalized problem and several variants of it can be written down in terms of fundamental vectors. We have developed here a simple and efficient algorithm to compute these fundamental vectors. Even in the non-normal case, this algorithm requires only order σ^2 operations to find Padé-Hermite forms of order σ. We remark also that it can be easily parallelized and/or vectorized. Working through an example, we have shown how simple it is to solve related problems which impose some extra conditions on the Padé-Hermite forms. In this paper, we followed a "diagonal" path through the Padé-Hermite table. In a forthcoming paper, we shall indicate that we can use a similar algorithm to go along any path we like in the Padé-Hermite table, even if it is a non-perfect one.

References

[1] *B. Beckermann.* Zur Interpolation mit polynomialen Linearkombinationen beliebiger Funktionen. *PhD thesis, Universität Hannover, 1989.*

[2] *B. Beckermann. A reliable method for computing M-Padé approximants on arbitrary staircases.* J. Comp. Appl. Math.*, 1990. Preprint.*

[3] *B. Beckermann. The structure of the singular solution table of the M-Padé approximation problem.* J. Comp. Appl. Math.*, 32(1 & 2):3–15, 1990.*

[4] *P.B. Borwein. Quadratic Hermite-Padé approximation to the exponential function.* Constr. Approx.*, 2:291–302, 1986.*

[5] *P.B. Borwein. Quadratic and higher order Padé approximants. In J. Szabados and K. Tandori, editors,* Alfred Haar Memorial Conference, Budapest, Hungary, 1985*, volume 49 of* Colloquia Mathematica Societatis János Bolyai*, pages 213–224. North-Holland, 1987.*

[6] *A. Bruckstein and T. Kailath. An inverse scattering framework for several problems in signal processing.* IEEE Acoust. Speech Signal Process. Magazine*, 4(1):6–20, 1987.*

[7] *A. Bultheel and M. Van Barel. Minimal vector Padé approximation.* J. Comp. Appl. Math.*, 32(1 & 2):27–37, 1990.*

[8] *G.V. Choodnovsky. Approximants rationelles des logarithmes des nombres rationnels.* C.R. Acad. Sc. Paris, Series A*, 288:607–609, 1979.*

[9] *G.V. Choodnovsky. Formules d'Hermite pour les approximants de Padé de logarithmes et de fontions binômes et mesures d'irrationalié.* C.R. Acad. Sc. Paris, Series A*, 288:965–967, 1979.*

[10] *G.V. Chudnovsky. Padé-approximations to the generalized hypergeometric functions i.* J. Math. Pures et Appl.*, 58:445–476, 1979.*

[11] *G.V. Chudnovsky. Un système explicité d'approximants de Padé pour des fonctions hypergéométriques généralisées avec applications à l'arithmétique.* C.R. Acad. Sc. Paris, Series A*, 288:1001–1003, 1979.*

[12] *J. Coates. On the algebraic approximation of functions, i-iii.* Proc. Kon. Nederl. Akad. Wet., *69:421–461, 1966. (Indagationes Mathematicae, 28).*

[13] *M.G. de Bruin.* Generalized continued fractions and a multidimensional Padé table. *PhD thesis, Amsterdam, 1974.*

[14] *M.G. de Bruin. Convergence in the Padé table for* ${}_1F_1(1, c; x)$. Indag. math., *38:408–418, 1976.*

[15] *M.G. de Bruin. Three new examples of generalized Padé tables which are partly normal. Technical Report 76-11, Dept. of Math., Univ. of Amsterdam, Amsterdam, 1976.*

[16] *M.G. de Bruin. Some convergence results in simultaneous rational approximation to the set of hypergeometric functions* $\{{}_1F_1(1; c_i, z)\}_{i=1}^n$. *In H. Werner and H.J. Bünger, editors,* Padé approximation and its applications, Bad Honnef 1983, *volume 1071 of* Lecture Notes in Mathematics, *pages 95–117, Berlin, 1983. Springer-Verlag.*

[17] *J. Della Dora.* Contribution à l'approximation de fonctions de la variable complex au sens de Hermite-Padé et de Hardy. *PhD thesis, l'Université Scientifique et Médicale de Grenoble, june 1980.*

[18] *J. Della Dora. Quelques resultats sur la structure des tables de Padé-Hermite. In M.G. de Bruin and H. van Rossum, editors,* Padé Approximation and its Applications, Lecture Notes in Mathematics 888, *pages 173–185, Berlin, 1980. Springer-Verlag.*

[19] *J. Della Dora and C. di Crescenzo. Approximants de Padé-Hermite, 1ère partie: théorie.* Numer. Math., *43:23–29, 1984.*

[20] *J. Della Dora and C. di Crescenzo. Approximants de Padé-Hermite, 2ème partie: programmation.* Numer. Math., *43:41–57, 1984.*

[21] *C. Hermite. Sur la fonction exponentielle. In* Œuvres, vol. III, *pages 150–181. Gauthier-Villars, Paris, 1912. Original = Compt. Rend. Acad. Sc., 77: 18-24, 226-233, 285-293, 1873.*

[22] *C. Hermite. Sur l'Expression* $U \sin x + C \cos x + W$. *Extrait d'une lettre à M. Paul Gordan. In* Œuvres, vol. III, *pages 135–145. Gauthier-Villars, Paris, 1912. Original = J. de Crelle, 76: 303-312, 1873.*

[23] *C. Hermite. Sur quelques approximations algébriques. Extrait d'une lettre à M. Borchardt. In* Œuvres, vol. III, *pages 146–149. Gauthier-Villars, Paris, 1912. Original = Journal de Crelle, 76: 342-344, 1873.*

[24] *C. Hermite. Sur la généralisation des fractions continues algébriques. Extrait d'une lettre à M. Pincherle. In* Œuvres, vol. IV, *pages 357–377. Gauthier-Villars, Paris, 1917. Original = Annali di Mathematica, Serie 2, 21: 289-108, 1893.*

[25] *H. Jager. A multidimensional generalization of the Padé table.* Proc. Kon. Nederl. Akad. Wet., *67:193–249, 1964. (Indagationes Mathematicae, 26).*

[26] *G. Labahn and S. Cabay. Matrix Padé fractions.* J. Comp. and Appl. Math., *1990. Preprint.*

[27] *W. Lübbe.* Über ein allgemeines Interpolationsproblem- Lineare Identitäten zwischen benachbarten Lösungssystemen. *PhD thesis, Universität Hannover, 1983.*

[28] *K. Mahler. Ein Beweis des Thue-Siegelschen Satzes über die Approximation algebraischen Zahlen für binomische Gleichungen.* Math. Ann.*, 105:267–276, 1931.*

[29] *K. Mahler. Zur Approximation der Exponential Funktion und des Logarithmus, Teil I.* Journal f. d. r. u. a.*, 166:118–136, 1932.*

[30] *K. Mahler. On the approximation of logarithms of algebraic numbers.* Philosophical transactions of the Royal Society of London, Series A*, 245:371–398, 1953.*

[31] *K. Mahler. Perfect systems.* Compositio Math.*, 19:95–166, 1968.*

[32] *E.M. Nikišin. On simultaneous Padé approximants.* Math. USSR Sbornik*, 41:409–425, 1982.*

[33] *H. Padé.* Sur la Représentation Approchée d'une Fonction par des Fractions Rationelles. *PhD thesis, Ann. Ecole. Norm. Sup., Paris, 1892.*

[34] *H. Padé.* Sur la Représentation Approchée d'une Fonction par des Fractions Rationelles. Supplement Thesis. *PhD thesis, Ann. Ecole. Norm. Sup., Paris, 1892.*

[35] *H. Padé. Sur la généralisation des fractions continues algébriques.* J. Math. Pures Appl., sér. 4*, 10:291–329, 1894.*

[36] *S. Paszkowski. Quelques algorithmes de l'approximation de Padé-Hermite. Technical Report Publication ANO 89, Univ. Lille I, 1982.*

[37] *S. Paszkowski. Quelques généralisations de l'approximation de Padé. Technical Report Publication ANO 93, Univ. Lille I, 1982.*

[38] *S. Paszkowski. Recurrence relations in Padé-Hermite approximation.* J. Comp. Appl. Math.*, 19(1):99–107, 1987.*

[39] *S. Paszkowski. Hermite-Padé approximation: basic notions and theorems.* J. Comp. Appl. Math.*, 32(1 & 2):229–236, 1990.*

[40] *R.E. Shafer. On quadratic approximation.* J. Numer. Anal.*, 11:447–460, 1974.*

[41] *M. Van Barel and A. Bultheel. A new approach to the rational interpolation problem.* J. Comp. Appl. Math.*, 32(1 & 2):281–289, 1990.*

[42] *M. Van Barel and A. Bultheel. A new approach to the rational interpolation problem: the vector case.* J. Comp. Appl. Math.*, 33(3):331–346, 1990.*